

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 21:44:31 / Search time 43 Seconds
(without alignments)
693.656 Million cell updates/sec

Title: US-10-736-227-1

Perfect score: 1663
Sequence: 1 MGNOVKETLHSLYKEVPTAD.....PPGRPPAPSSSEEDGEAAAH 310

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: PIR_80: +
1: p1r1: +
2: p1r2: +
3: p1r3: +
4: p1r4: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Q. suIt No.	Score	Query Match	Length	DB ID	Description
1	116	7.0	160	2 S42794	p18H-rev 107 prote
2	102.5	6.2	249	2 A41497	36K antigen pra -
3	102	6.1	1956	2 T00051	hypothetical prote
4	101.5	6.1	673	2 T00328	hypothetical prote
5	97.5	5.9	419	2 T29286	hypothetical prote
6	97.5	5.9	516	2 H71332	hypothetical prote
7	97.5	5.9	775	2 S55345	protein-tyrosine-P
8	96.5	5.8	1206	2 S24407	protein isoform IV
9	96.5	5.8	1469	2 S11515	formin - mouse
10	95.5	5.7	189	2 AE3327	hypothetical prote
11	95.5	5.7	301	2 E29149	proline-rich prote
12	95.5	5.7	549	2 C87719	protein R119.6 (Im
13	94.5	5.7	481	2 T27975	hypothetical prote
14	94.5	5.7	500	2 S64220	hypothetical prote
15	94	5.7	240	2 D70894	probable pra prote
16	93.5	5.6	229	1 W4W18	E4 protein - human
17	93.5	5.6	416	1 W2W1DP	E2 protein - deer
18	93.5	5.6	465	1 A33647	sulfated surface g
19	93	5.6	245	1 W4W1S	E4 protein - human
20	92.5	5.6	1952	2 T48814	hypothetical prote
21	92	5.5	171	2 A27307	proline-rich phosp
22	92	5.5	472	1 A49836	transcripton fact
23	92	5.5	636	2 T07447	DNA-directed RNA p
24	91.5	5.5	281	2 S65692	E4 protein - human
25	91.5	5.5	508	2 S54264	glycoprotein GC -
26	91.5	5.5	2957	2 T33152	hypothetical prote
27	91	5.5	385	2 H70937	hypothetical prote
28	91	5.5	1239	1 QOBS10	BOLFI protein - hu
29	91	5.5	1692	2 G01449	probable mucin G2

30	90.5	5.4	318	2 T29479	hypothetical prote
31	90.5	5.4	1392	2 T51947	probable transcript
32	90.5	5.4	2715	2 T13049	eyelid - fruit fly
33	90	5.4	372	2 T14193	L-ascorbate peroxi
34	90	5.4	521	2 A29249	RNA-directed RNA p
35	90	5.4	741	2 I51657	suppressor of yeas
36	89.5	5.4	170	2 A48013	proline-rich prote
37	89.5	5.4	286	1 P1RT3	acidic proline-ric
38	89.5	5.4	204	2 H87524	hypothetical prote
39	89.5	5.4	420	2 T46910	hypothetical prote
40	89.5	5.4	521	1 VGBEBH	glycoprotein gill
41	89.5	5.4	697	1 TVPPTL	large T antigen -
42	89.5	5.4	1018	2 T43168	hypothetical prote
43	89.5	5.4	1029	1 S21369	collagen alpha 2IV
44	89.5	5.4	1259	2 T16038	hypothetical prote
45	89.5	5.4	1638	2 T30313	chemotaxis protein

ALIGNMENTS

RESULT 1

S42794
p18H-rev 107 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S42794
R:Schaefer, R.
submitted to the EMBL Data Library, November 1993
A:Reference number: S42794
A:Accession: S42794
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <SCH>
A:Cross-references: UNIPROT:P53817; UNIPARC:UPI000012CEB5; EMBL:X76453; NID:G433962; PI.

Query Match 7.0%; Score 116; DB 2; Length 160;
Best Local Similarity 31.9%; Pred. No. 0.024;
Matches 38; Conservative 17; Mismatches 36; Indels 28; Gaps 6;

Qy 117 KPGLVFEVSOQYPPHNAVYGNFQVHL-----HRLVINSFLTDS----- 159
Db 7 KPGLIE-IFRPMYSHAIYVGDDYVHLAPPSEIPGAGASWLSALTDAIVKLELRD 65
Qy 160 -QGRGRVNDLY--RYKPLSSAVVNNALAHVGAKEKRELSWR-----NSESFAWCRYG 211
Db 66 VAGKDKYQVNNKDKETTPLPNKTIQRAEELVG---QEVYRUTSNCERFVNELYRG 121

RESULT 2

A41497
36K antigen pra - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A41497
R:Thole, J.E.R.; Stabel, L.F.E.M.; Suykerbuyk, M.E.G.; de Wit, M.Y.L.; Klatter, P.R.; K.
Infect. Immun. 58, 80-87, 1990
A:Title: A major immunogenic 36,000-molecular-weight antigen from Mycobacterium leprae
A:Reference number: A41497; PMID:90093489; PMID:1688422
A:Accession: A41497
A:Molecule type: DNA
A:Residues: 1-249 <THO>
A:Cross-references: UNIPROT:P41484; UNIPARC:UPI000013216A; GB:X65546

Query Match 6.2%; Score 102.5; DB 2; Length 249;
Best Local Similarity 20.4%; Pred. No. 0.54;
Matches 56; Conservative 26; Mismatches 108; Indels 85; Gaps 11;

Qy 17 PTADPTGVDRDDPRIGSVYIFSNDDVDVPPPPGPGGAGL-----PDGADGPPPPPP 71
Db 14 PAPPFGSSGSGYRPPSPFSELGS-----AYPPFPYVGGSYPPPPPGGSGYPPPPPP 66
Qy 72 QPYDPRLHEVEGVSFYRDECTYOKSFAPGSAALSTYTPENLLNKKCRGLDIVERVSOQVYP 131

Db 301 EEDGEAVAH 310

RESULT 2

US-10-074-475-267
; Sequence 267, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Kaira, Kaipana
; APPLICANT: Cafferty, Robert
; APPLICANT: Sun, Yonnging
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 267
; LENGTH: 111
; TYPE: PRN
; ORGANISM: Homo sapien
US-10-074-475-267

Query Match 27.0%; Score 449.5; DB 4; Length 111;
Best Local Similarity 82.9%; Pred. No. 1.6e-32;
Matches 92; Conservative 4; Mismatches 14; Indels 1; Gaps 1;

Qy 168 NDLYRYKPLSSAVYRNALAHVAKRELSWRNSESPPAWGCKREKIGGELRIGKOP 227
Db 1 NDLYRYKPLSSAVYRNALAHVAKRELSWRNSESPPAWGCKREKIGGELRIGKOP 60
228 YRLQIQLSAQRSHLTLEFQSLLEDLIM-EKRRNDQIGRAVLOELATHTLHPAE 277
61 YRLQIQLSAQRSHLTLEFQSLLEDLIMGEATQPRSGARPVLOELATHTLHPAE 111

RESULT 3

US-10-104-047-3151
; Sequence 3151, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3151
; LENGTH: 354
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-104-047-3151

Query Match 16.2%; Score 269.5; DB 4; Length 354;
Best Local Similarity 32.8%; Pred. No. 1.5e-15;
Matches 76; Conservative 25; Mismatches 80; Indels 51; Gaps 10;

Qy 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSYIFSNDDVEVEPPQPGDGGSLP 60
Db 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSYIFSNDDVEVEPPQPGDGGSLP 60
61 D-GGDGPPPPQPPQPPYDRLH-----EVECSVFYRDECIYOK-SFAPGSAALST 106

Db 51 DKRGVKAPEGCTPCPESPSRHHHLLHQLVLTETQFSAFQCECTFSKVSQPGADLSV 110

Qy 107 YTPENLTKCKPGDLVEFV-----SOAQYPMWVYVGNFQVHLHRELVINSFLTDAS 159

Db 111 YAVTALPACPEGDLLELWLOPAPPEPPAPAPHWASPAR-SAGRTRRASPPGAAALASGS 169

Qy 160 QGRGRVNDLYRYKPLSSAVYRNALAHVAKRELSWRNSES-----FAAW 207
Db 170 SRREGRC--RQARSPRSSSTISR-----CTWERTRSTPPGFTAW 206

RESULT 4

US-11-072-512-3151
; Sequence 3151, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHINO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3151
; LENGTH: 354
; TYPE: PRN
; ORGANISM: Homo sapiens
US-11-072-512-3151

Query Match 16.2%; Score 269.5; DB 6; Length 354;
Best Local Similarity 32.8%; Pred. No. 1.5e-15;
Matches 76; Conservative 25; Mismatches 80; Indels 51; Gaps 10;

Qy 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSYIFSNDDVEVEPPQPGDGGSLP 60
Db 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSYIFSNDDVEVEPPQPGDGGSLP 60

Qy 61 D-GGDGPPPPQPPQPPYDRLH-----EVECSVFYRDECIYOK-SFAPGSAALST 106
Db 51 DKRGVKAPEGCTPCPESPSRHHHLLHQLVLTETQFSAFQCECTFSKVSQPGADLSV 110

Qy 107 YTPENLTKCKPGDLVEFV-----SOAQYPMWVYVGNFQVHLHRELVINSFLTDAS 159
Db 111 YAVTALPACPEGDLLELWLOPAPPEPPAPAPHWASPAR-SAGRTRRASPPGAAALASGS 169

Qy 160 QGRGRVNDLYRYKPLSSAVYRNALAHVAKRELSWRNSES-----FAAW 207
Db 170 SRREGRC--RQARSPRSSSTISR-----CTWERTRSTPPGFTAW 206

RESULT 5

```

US-10-450-763-33497
; Sequence 33497, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 33497
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(96)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-33497

Query Match      13.8%; Score 229; DB 5; Length 96;
Best Local Similarity 66.2%; Pred. No. 1.2e-12;
Matches 45; Conservative 3; Mismatches 20; Indels 0; Gaps 0

QY      52  GGPDSGGLPDGGDDGPFPQPCYDPDRHLHVEGCVYPRDCTCYQKSPAPGSAALSTYTPEN 111
Db      11  RGNMAAPACTYGTGCRPRRSRSYDPRHLHVECVYPRDCTCYQKSPAPGSAALSTYTPEN 70
QY      112 LLNKCKPG 119
Db      71  LPQGVQAG 78

RESULT 6
US-10-171-311-198
; Sequence 198, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoeseish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRL-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-198

```

```

Query Match          8.5%; Score 141; DB 4; Length 164;
Best Local Similarity 32.2%; Pred. No. 0.00021;
Matches 39; Conservative 21; Mismatches 39; Indels 22; Gaps 5;

OY      117 KPGDLVEFVSQAQYPHWATVGNFOVHL-----HRLVINSLFTDASGCRG---- 164
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       9 KPGLDLE-IFRLGEYHMAIYIGGYVIHLAPSEYPCAGSSVFYSLSNSAEVKRGRLED 67
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

OY      165 -----RVVNDL-YRYKPLSSSAVVVNALAHGAKER-ELSRNSESFAAMCRYGKREF 215
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       68 VVGCCCRVNNSLDHETQRPPEVITISSAKENVGMKKTISVRNCBHFVAQLRTGKSRC 127
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

OY      216 K 216
        |
DB       128 K 128

RESULT 7
US-10-723-860-3892
Sequence 3892, Application US/10723860
Publication No. US20040253606A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Ginsburg, Wendy M.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
FILE REFERENCE: 05882.0193.NPUS01
CURRENT APPLICATION NUMBER: US/10/723,860
CURRENT FILING DATE: 2003-11-26
PRIOR APPLICATION NUMBER: 60/429,739
PRIOR FILING DATE: 2002-11-26
NUMBER OF SEQ ID NOS: 8393
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3892
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
US-10-723-860-3892

Query Match          8.5%; Score 141; DB 5; Length 164;
Best Local Similarity 32.2%; Pred. No. 0.00021;
Matches 39; Conservative 21; Mismatches 39; Indels 22; Gaps 5;

OY      117 KPGDLVEFVSQAQYPHWATVGNFOVHL-----HRLVINSLFTDASGCRG---- 164
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       9 KPGLDLE-IFRLGEYHMAIYIGGYVIHLAPSEYPCAGSSVFYSLSNSAEVKRGRLED 67
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

OY      165 -----RVVNDL-YRYKPLSSSAVVVNALAHGAKER-ELSRNSESFAAMCRYGKREF 215
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB       68 VVGCCCRVNNSLDHETQRPPEVITISSAKENVGMKKTISVRNCBHFVAQLRTGKSRC 127
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

OY      216 K 216
        |
DB       128 K 128

RESULT 8
US-10-756-149-5683
Sequence 5683, Application US/10756149
Publication No. US20050181375A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS &
FILE REFERENCE: file
CURRENT APPLICATION NUMBER: US/10/756,149
CURRENT FILING DATE: 2004-01-12
NUMBER OF SEQ ID NOS: 5818
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5683
LENGTH: 164

```

Db 9 KPQDILIE-I FRLGYEHWALYIGDGYIHLAPSEYPGAGSSSVFSLSNSAEVKKRRLD 67

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 16
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
US-10-616-942-16

Query Match
Best Local Similarity 32.2%; Score 136; DB 4; Length 164;
Matches 39; Conservative 15; Mismatches 45; Indels 22; Gaps 4;

QY 117 KPGDVEFVSQAQYPMHAYVGNFOYVHL-----HLEVINSLTD 157
DB 9 KPGDIE-IFRLGYHEHMAIYIDGVYIHLAPPESEYPCAGSSVSFVLSNSAEVKEERLTD 67
QY 158 ASQGRGRVNDL-TRYKPLSSAVVRNALAHVAKER-ELSMRNSSEFAACRCYGRPF 215
DB 68 VVGCCYRVNNSLDHEYOPRVEVITISSAKENVGQMKYSIVSRNCEHFVTLRYGKSRG 127
QY 216 K 216
DB 128 K 128

RESULT 12
US-10-116-326-6

Sequence 6, Application US/10116326
Publication No. US20030166889A1

GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Mathur, Brian

APPLICANT: Fiedler, Carl Johan

TITLE OF INVENTION: No. US20030166889A1 Human Kinases and Polynucleotides Encoding

FILE REFERENCE: LEX-0332-USA

CURRENT APPLICATION NUMBER: US/10/116,326

CURRENT FILING DATE: 2002-04-04

PRIOR APPLICATION NUMBER: US 60/282,036

PRIOR FILING DATE: 2001-04-06

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 703

TYPE: PRT

ORGANISM: homo sapiens

US-10-116-326-6

Query Match
Best Local Similarity 21.6%; Score 122; DB 4; Length 703;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

QY 50 PPOGPDGGGLPDGDPGPPPOQ-----POPYDP----- 76
DB 404 PSRGRGGG---AGQPPPPSARSTPLPQPPSPSSGGTPLHSPHTPRASPTGTGTT 460
QY 77 -----RLHEVECSVYRDECIYKSFAGSALSTYTPENLKKCKPD 120
DB 461 PPPSPGGVGGAAMSRINSFLGSPRFRHRRKQVTAEMSLTPRESSPELAKRS- 519
QY 121 LVEFVSQAQYPMHAYVGNFOYVHLHLEVINSLTDSQGRGRVNDLYRYKPLSSA 180
DB 520 -----WFGNF--ISLDKEQITLVLDK-----KPLSS-- 544
QY 181 VVRNALAHVAKERELSWR--NSSEFAACRCYGRKEFGELRIQKOPRYLQIOLSAQR 238
DB 545 -IKADIVAFSLPSLSHVSLSQTSFRA-----EYKASGSPSVFOKRVKQVVISSE 596
QY 239 S-----HTLEFQSLDLMKRRNDQIGRAAVLQELATLHHPAEEED- 282
DB 597 GPEPSRRDGGGGIYGVTFLLISG---PSRRFRKVVETIOAQLSLTHDPSVOLADE 653
QY 283 SNVART-----TPPGRP--PAPSEEDG 305
DB 654 KNGAOTRPAAGAPPRSLQPPGGRPDELSSSPRRG 687

RESULT 13
US-10-803-277-6

Sequence 6, Application US/10803277
Publication No. US20040180416A1

GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Mathur, Brian

APPLICANT: Fiedler, Carl Johan

TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same

FILE REFERENCE: LEX-0332-USA

CURRENT APPLICATION NUMBER: US/10/803,277

CURRENT FILING DATE: 2004-03-18

PRIOR APPLICATION NUMBER: US/10/116,326

PRIOR FILING DATE: 2002-04-04

PRIOR APPLICATION NUMBER: US 60/282,036

PRIOR FILING DATE: 2001-04-06

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 703

TYPE: PRT

ORGANISM: homo sapiens

US-10-803-277-6

Query Match
Best Local Similarity 21.6%; Score 122; DB 4; Length 703;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

QY 50 PPOGPDGGGLPDGDPGPPPOQ-----POPYDP----- 76
DB 404 PSRGRGGG---AGQPPPPSARSTPLPQPPSPSSGGTPLHSPHTPRASPTGTGTT 460
QY 77 -----RLHEVECSVYRDECIYKSFAGSALSTYTPENLKKCKPD 120
DB 461 PPPSPGGVGGAAMSRINSFLGSPRFRHRRKQVTAEMSLTPRESSPELAKRS- 519
QY 121 LVEFVSQAQYPMHAYVGNFOYVHLHLEVINSLTDSQGRGRVNDLYRYKPLSSA 180
DB 520 -----WFGNF--ISLDKEQITLVLDK-----KPLSS-- 544
QY 181 VVRNALAHVAKERELSWR--NSSEFAACRCYGRKEFGELRIQKOPRYLQIOLSAQR 238
DB 545 -IKADIVAFSLPSLSHVSLSQTSFRA-----EYKASGSPSVFOKRVKQVVISSE 596
QY 239 S-----HTLEFQSLDLMKRRNDQIGRAAVLQELATLHHPAEEED- 282
DB 597 GPEPSRRDGGGGIYGVTFLLISG---PSRRFRKVVETIOAQLSLTHDPSVOLADE 653
QY 283 SNVART-----TPPGRP--PAPSEEDG 305
DB 654 KNGAOTRPAAGAPPRSLQPPGGRPDELSSSPRRG 687

RESULT 14
US-10-204-639-65
Sequence 65, Application US/10204639
Publication No. US20060063152A1

GENERAL INFORMATION:
APPLICANT: Osamu Ohara
APPLICANT: Takahiro Nagase
APPLICANT: Daisuke Nakajima
TITLE OF INVENTION: NOVEL GENE AND PROTEIN ENCODED BY THE GENE
FILE REFERENCE: PH-1416-PCT
CURRENT APPLICATION NUMBER: US/10/204,639
CURRENT FILING DATE: 2002-08-22
PRIOR APPLICATION NUMBER: JP 2000-389742
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: JP 2001-095524
PRIOR FILING DATE: 2001-03-29
PRIOR APPLICATION NUMBER: JP 2001-127066
PRIOR FILING DATE: 2001-04-25
NUMBER OF SEQ ID NOS: 140

SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-204-639-65

Query Match 7.3%; Score 122; DB 5; Length 715;
Best Local Similarity 21.6%; Pred. No. 0.082;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

QY 50 PPGPGGCGLPDGGDGGPPPPQ-----PPPYDP----- 76
DB 416 PSRGPGRGG--AGEQPPPSARSTPLPGPGSPRSSGGTPLHSPLHTPRASPTGTGCTT 472
QY 77 -----RLHEVECSVFYRDECIYOKSFAPGSALSTYTPENLANKCKPGD 120
DB 473 PPSPGCGVGGAAMRSRLNSFLGSPFRHRKMQVPTAEEMSLTPSSPELAKRS- 531
QY 121 LVEFVSQAQYPHMAVYVGNFQVHLHRLLEVINSFLTASQGRGRVNDLRYRKLSSSA 180
3 532 -----WFGNF--ISLDKEQIFLVVKD-----KPLSS-- 556
2Y 181 VARNALAHGAKRELSMR--NSESFAAMCRYGKREFFKIGGELRIKOPYRLOIQLSAQR 238
DB 557 -IKADIVHAFSLIPSLSHSVLSQTSFRA-----ETKASGGSVPFQKPVRFQVDISSSE 608
QY 239 S-----HTLEFOSLEDLIMEKRRNDQIGRAAVLQELATHLHPAPEEGD- 282
DB 609 GPEPSRRDSSGGGCIYSVFTLISG--PSRRFKRVETIQQLSTHDQPSVQALADE 665
QY 283 SNVART-----TPPGRP-PAPSSEEDG 305
DB 666 KNGAQTRPAGAPPRLQPPGRDPPELSSSPRRG 699

RESULT 15

US-10-116-326-4
; Sequence 4, Application US/10116326
; Publication No. US2003016689A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US2003016689A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0332-USA
; CURRENT APPLICATION NUMBER: US/10/116,326
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,036
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 762
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-116-326-4

Query Match 7.3%; Score 122; DB 4; Length 762;
Best Local Similarity 21.6%; Pred. No. 0.09;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

QY 50 PPGPGGCGLPDGGDGGPPPPQ-----PPPYDP----- 76
DB 463 PSRGPGRGG--AGEQPPPSARSTPLPGPGSPRSSGGTPLHSPLHTPRASPTGTGCTT 519
QY 77 -----RLHEVECSVFYRDECIYOKSFAPGSALSTYTPENLANKCKPGD 120
DB 520 PPSPGCGVGGAAMRSRLNSFLGSPFRHRKMQVPTAEEMSLTPSSPELAKRS- 578
QY 121 LVEFVSQAQYPHMAVYVGNFQVHLHRLLEVINSFLTASQGRGRVNDLRYRKLSSSA 180
DB 579 -----WFGNF--ISLDKEQIFLVVKD-----KPLSS-- 603

QY 181 VARNALAHGAKRELSMR--NSESFAAMCRYGKREFFKIGGELRIKOPYRLOIQLSAQR 238
DB 604 -IKADIVHAFSLIPSLSHSVLSQTSFRA-----ETKASGGSVPFQKPVRFQVDISSSE 655
QY 239 S-----HTLEFOSLEDLIMEKRRNDQIGRAAVLQELATHLHPAPEEGD- 282
DB 656 GPEPSRRDSSGGGCIYSVFTLISG--PSRRFKRVETIQQLSTHDQPSVQALADE 712
QY 283 SNVART-----TPPGRP-PAPSSEEDG 305
DB 713 KNGAQTRPAGAPPRLQPPGRDPPELSSSPRRG 746

Search completed: December 6, 2006, 21:54:21
Job time : 189 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 21:51:21 : Search time 40 Seconds

(without alignments)
699.504 Million cell updates/sec

Title: US-10-736-227-1

Perfect score: 1663
Sequence: 1 MGNVEKTLTHSYKEVPTAD.....PGRPPAPSSSEEDGEVAH 310

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 358988 seqs, 90258633 residues

Total number of hits satisfying chosen parameters: 358988

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US09_NEW_PUB pep:*
- 2: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US06_NEW_PUB pep:*
- 3: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US07_NEW_PUB pep:*
- 4: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US08_NEW_PUB pep:*
- 5: /EMC_Celerra_SIDS3/prodata/1/pubpaa/PCT_NEW_PUB pep:*
- 6: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US10_NEW_PUB pep:*
- 7: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US11_NEW_PUB pep:*
- 8: /EMC_Celerra_SIDS3/prodata/1/pubpaa/US60_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1663	100.0	310	US-10-527-788-63	Sequence 63, Appl
2	642.5	38.6	292	US-10-527-788-64	Sequence 64, Appl
3	282	17.0	63	US-10-527-788-23	Sequence 23, Appl
4	282	17.0	63	US-10-527-788-34	Sequence 34, Appl
5	141	8.5	164	US-10-533-519-822	Sequence 822, Appl
6	122	7.3	794	US-11-320-072-74	Sequence 74, Appl
7	115.5	6.9	872	US-10-613-783-4430	Sequence 4430, Appl
8	101.5	6.1	328	US-10-131-833A-364	Sequence 364, Appl
9	97.5	5.9	854	US-10-449-902-41174	Sequence 41174, Appl
10	94.5	5.7	543	US-10-953-349-21087	Sequence 21087, Appl
11	94.5	5.7	586	US-10-953-349-21086	Sequence 21086, Appl
12	94.5	5.7	1453	US-11-056-355B-81762	Sequence 81762, Appl
13	94.5	5.7	1661	US-11-056-355B-81761	Sequence 81761, Appl
14	94.5	5.7	1714	US-11-056-355B-81760	Sequence 81760, Appl
15	93.5	5.6	326	US-11-056-355B-1197	Sequence 1197, Appl
16	93.5	5.6	326	US-10-449-902-40809	Sequence 40809, Appl
17	93	5.6	1081	US-10-612-783-6929	Sequence 6929, Appl
18	92.5	5.6	270	US-10-953-349-26560	Sequence 26560, Appl
19	92.5	5.6	270	US-11-056-355B-63759	Sequence 63759, Appl
20	92.5	5.6	309	US-10-953-349-26559	Sequence 26559, Appl
21	92.5	5.6	309	US-11-056-355B-63758	Sequence 63758, Appl
22	92.5	5.6	357	US-10-449-902-50208	Sequence 50208, Appl
23	92.5	5.6	447	US-10-539-228-684	Sequence 684, Appl
24	92	5.5	567	US-10-449-902-45116	Sequence 45116, Appl
25	92	5.5	1015	US-10-553-520-132	Sequence 132, Appl

26	91.5	5.5	773	US-10-539-228-845	Sequence 845, Appl
27	91	5.5	597	US-11-043-842-765	Sequence 765, Appl
28	91	5.5	597	US-11-362-260A-173	Sequence 173, Appl
29	90.5	5.4	849	US-10-953-349-1431	Sequence 1431, Appl
30	90.5	5.4	1003	US-10-953-349-1431	Sequence 1431, Appl
31	90.5	5.4	1392	US-10-953-349-1430	Sequence 1430, Appl
32	90.5	5.4	1616	US-10-533-519-52	Sequence 52, Appl
33	90	5.4	346	US-10-953-349-14821	Sequence 14821, Appl
34	90	5.4	346	US-11-056-355B-1723	Sequence 4723, Appl
35	90	5.4	347	US-11-056-355B-73950	Sequence 73950, Appl
36	90	5.4	348	US-10-953-349-6363	Sequence 6363, Appl
37	90	5.4	348	US-11-056-355B-25741	Sequence 25741, Appl
38	90	5.4	348	US-11-056-355B-35941	Sequence 35941, Appl
39	90	5.4	348	US-11-056-355B-105997	Sequence 105997, Appl
40	90	5.4	348	US-11-056-355B-117236	Sequence 117236, Appl
41	90	5.4	352	US-11-056-355B-73949	Sequence 73949, Appl
42	90	5.4	353	US-11-056-355B-25740	Sequence 25740, Appl
43	90	5.4	372	US-10-953-349-6362	Sequence 6362, Appl
44	90	5.4	372	US-11-056-355B-35940	Sequence 35940, Appl
45	90	5.4	372	US-11-056-355B-105996	Sequence 105996, Appl

ALIGNMENTS

RESULT 1
US-10-527-788-63
Sequence 63, Application US/10527788

Publication No. US20060172300A1

GENERAL INFORMATION:

US-10-527-788-63

Query Match 100.0%; Score 1663; DB 6; Length 310;

Best Local Similarity 100.0%; Pred. No. 1.5e-125;

Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGNVEKTLTHSYKEVPTADPTGVRDDGPRIGSYFNSNDEDEVEPPPGGGLP	60
DB	1	MGNVEKTLTHSYKEVPTADPTGVRDDGPRIGSYFNSNDEDEVEPPPGGGLP	60
QY	61	DGGGPPPPQPPDPRLHEVEGVFPRDECIYKSFAPGSAALSTYPPENLNKCKPGD	120
DB	61	DGGGPPPPQPPDPRLHEVEGVFPRDECIYKSFAPGSAALSTYPPENLNKCKPGD	120
QY	121	LVEFVSQAQYDHMAVYVGNFQVHLHRLVINSFLTJASQGRGRVYNDLYRYKPLSSA	180
DB	121	LVEFVSQAQYDHMAVYVGNFQVHLHRLVINSFLTJASQGRGRVYNDLYRYKPLSSA	180
QY	181	VVRNALAHVGAKEKELSWRNESFPAMCRVYKREFKIGELRIGKOPRIQIOLSAQRSH	240
DB	181	VVRNALAHVGAKEKELSWRNESFPAMCRVYKREFKIGELRIGKOPRIQIOLSAQRSH	240

QY 241 TLEFOSIEDLIMEKRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSS 300
DB 241 TLEFOSIEDLIMEKRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSS 300
QY 301 EEDDEGAVAH 310
DB 301 EEDDEGAVAH 310

RESULT 2

US-10-527-788-64
Sequence 64, Application US/10527788
Publication No. US20060172300A1
GENERAL INFORMATION:
APPLICANT: WOLF, SABINE
APPLICANT: JAGER, MARTINA
APPLICANT: BANGSOM, THORSTEN
APPLICANT: BANGSOM, CARMEN
APPLICANT: JORDAN, DOMINIK
APPLICANT: PELZER, BERNHARD
APPLICANT: OPOLZER, THOMAS
TITLE OF INVENTION: METHOD FOR IDENTIFYING BBB-SPECIFIC PROTEINS AND
FILE REFERENCE: 63047(45107)
CURRENT APPLICATION NUMBER: US/10/527,788
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: PCT/EP03/09968
PRIOR FILING DATE: 2003-03-08
PRIOR APPLICATION NUMBER: DE 102 42 016.5
PRIOR FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 3.3
LENGTH: 292
SEQ ID NO 64
TYPE: PRT
ORGANISM: Homo sapiens
US-10-527-788-64

Query Match 38.6%; Score 642.5; DB 6; Length 292;
Best Local Similarity 46.4%; Pred. No. 7e-44;
Matches 137; Conservative 36; Mismatches 87; Indels 35; Gaps 7;
QY 1 MGNQVEGLTLSTYKEVPTADPTGVDRDDGRIGISYIFSNDDVEVEQPFGGDDGGLP 60
DB 1 MGNQVLDITLHNSYSELPTGDPGIGKDE-LRVGVAFSSDDEDL-----DERGQP 50
QY 61 D--GGDPPPPQPPQVPRLL-----EVECSVFYRDECIYOK-SFAPGSALST 106
DB 51 DKRGVAPRPGCTPCPESPSPRHQHLLHQLVLNETQFSAFGQECIFSKVSGPGADLSV 110
QY 107 YTPENLKKCKPGDLVEFV-----SOQOYPMNAVYVGNFQVVLHRLLEVINSPLTAS 159
DB 111 YATYALPALCEPGDLELMLQAPRPPAPAPAHVAVVGGQIHLHQGEIRQDSTLEAG 170
QY 160 QGRRGRVNDLYRKPLSSSAVYNALAHVAKRELSMRSEFPAMCRGKEFKIGG 219
DB 171 AAVNVRVNSWYRPLVAELVONACGHLGLKKEEICWNTSESFAAMCRGKKEFVAGG 230
QY 220 ELRIGKQP-----YRLOIOLSAORSHTEFOSLEDLIMEKRRNDQIGRAAVLOELA 270
DB 231 EVFAGTPOPOQOYLKVLHGENKVTARFHSLEDLIMEKRRIDASGLRVLOELA 285

RESULT 3

US-10-527-788-23
Sequence 23, Application US/10527788
Publication No. US20060172300A1
GENERAL INFORMATION:
APPLICANT: WOLF, SABINE
APPLICANT: JAGER, MARTINA
APPLICANT: BANGSOM, THORSTEN
APPLICANT: BANGSOM, CARMEN

APPLICANT: JORDAN, DOMINIK
APPLICANT: PELZER, BERNHARD
APPLICANT: OPOLZER, THOMAS
TITLE OF INVENTION: METHOD FOR IDENTIFYING BBB-SPECIFIC PROTEINS AND
FILE REFERENCE: 63047(45107)
CURRENT APPLICATION NUMBER: US/10/527,788
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: PCT/EP03/09968
PRIOR FILING DATE: 2003-03-08
PRIOR APPLICATION NUMBER: DE 102 42 016.5
PRIOR FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 23
LENGTH: 63
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Partial
amino acid sequence of NSE2 from swine
US-10-527-788-23

Query Match 17.0%; Score 282; DB 6; Length 63;
Best Local Similarity 85.7%; Pred. No. 7.9e-16;
Matches 54; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 248 EDLIMERRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSSSEEDGEA 307
DB 1 EDLIMERRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSSSEEDGEA 60

QY 308 VAH 310
DB 61 VVH 63

RESULT 4

US-10-527-788-34
Sequence 34, Application US/10527788
Publication No. US20060172300A1
GENERAL INFORMATION:
APPLICANT: WOLF, SABINE
APPLICANT: JAGER, MARTINA
APPLICANT: BANGSOM, THORSTEN
APPLICANT: BANGSOM, CARMEN
APPLICANT: JORDAN, DOMINIK
APPLICANT: PELZER, BERNHARD
APPLICANT: OPOLZER, THOMAS
TITLE OF INVENTION: METHOD FOR IDENTIFYING BBB-SPECIFIC PROTEINS AND
FILE REFERENCE: 63047(45107)
CURRENT APPLICATION NUMBER: US/10/527,788
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: PCT/EP03/09968
PRIOR FILING DATE: 2003-03-08
PRIOR APPLICATION NUMBER: DE 102 42 016.5
PRIOR FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 34
LENGTH: 63
TYPE: PRT
ORGANISM: Sus sp.
US-10-527-788-34

Query Match 17.0%; Score 282; DB 6; Length 63;
Best Local Similarity 85.7%; Pred. No. 7.9e-16;
Matches 54; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 248 EDLIMERRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSSSEEDGEA 307
DB 1 EDLIMERRRNDQIGRAAVLOELATHLHPAEPEDSDSVARTTTPPGRPAPSSSEEDGEA 60

QY	308	VAH	310
Db	61	VVH	63

RESULT 5
US-10-533-519-8222

```

Sequence 822, Application US/10533519
Publication No. US20060263774A1
GENERAL INFORMATION:
APPLICANT: CLARK, HILARY
APPLICANT: SCHOENFELD, JILL
APPLICANT: VANLOOKEREN, MENNO
APPLICANT: WILLIAMS, P. MICKEY
APPLICANT: WOOD, WILLIAM I.
APPLICANT: WU, THOMAS D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS OF IMMUNE
FILE REFERENCE: P1904R1 US
CURRENT APPLICATION NUMBER: US/10/533,519
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: PCT/US03/34312
PRIOR FILING DATE: 2003-10-30
PRIOR APPLICATION NUMBER: US 60/423,394
PRIOR FILING DATE: 2002-11-01
NUMBER OF SEQ ID NOS: 2517
SEQ ID NO 822
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapien
US-10-533-519-822

```

Query Match	8.5%;	Score 141;	DB 6;	Length 164
Best Local Similarity	32.2%;	Pred. No. 0.00049;		
Matches	39;	Conservative	21;	Mismatches 39;
				Indels

[illegible]

```

1      RESULT 6
2      US-11-320-072-74
3      Sequence 74, Application US/11320072
4      Publication No. US20060216730A1
5      GENERAL INFORMATION:
6      APPLICANT: PLOWMAN, GREGORY D.
7      APPLICANT: WHYTE, DAVID
8      APPLICANT: MANNING, GERRARD
9      APPLICANT: SUDARSSANAM, SUCHA
10     APPLICANT: MARTINEZ, RICARDO
11     APPLICANT: FLANAGAN, PETER
12     APPLICANT: CLARY, DOUGLAS
13     TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
14     TITLE OF INVENTION: ENZYMES
15     FILE REFERENCE: 034536-2157
16     CURRENT APPLICATION NUMBER: US/11/320,072
17     CURRENT FILING DATE: 2005-12-29
18     PRIOR APPLICATION NUMBER: 10/110,978
19     PRIOR FILING DATE: 2002-10-29
20     PRIOR APPLICATION NUMBER: PCT/US00/32085
21     PRIOR FILING DATE: 2000-11-22
22     PRIOR APPLICATION NUMBER: 60/1167,482
23     PRIOR FILING DATE: 1999-11-24
24     PRIOR APPLICATION NUMBER: 60/166,997

```

```

PRIORITY FILING DATE: 1999-12-03
PRIORITY APPLICATION NUMBER: 60/114,185
PRIORITY FILING DATE: 1999-12-27
PRIORITY APPLICATION NUMBER: 60/118,078
PRIORITY FILING DATE: 2000-01-25
PRIORITY APPLICATION NUMBER: 60/119,364
PRIORITY FILING DATE: 2000-01-31
PRIORITY APPLICATION NUMBER: 60/183,173
PRIORITY FILING DATE: 2000-02-17
PRIORITY APPLICATION NUMBER: 60/187,150
PRIORITY FILING DATE: 2000-03-06
PRIORITY APPLICATION NUMBER: 60/190,162
PRIORITY FILING DATE: 2000-03-17
Remaining Prior Application data removed - See File Wrapper or PALM
NUMBER OF SEQ ID NOS: 277
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 74
LENGTH: 794
TYPE: PRF
ORGANISM: Homo sapiens
US-11-320-072-74

```

Query Match	7.3%;	Score 122;	DB 7;	Length 794;
Best Local Similarity	21.6%;	Pred. No. 0.11;		
Matches	72;	Conservative	32;	Mismatches 102;
				Indels 128;
				Gaps 14;

```

Oy 50 PPGPGGGGLPGGSGPPPPQ---PQYDP-----76
      |::|||::|||
Db 495 PSRGPPRGGG---AGGQPPPPSARSTPLGPPGSPSSSGTFLHSPHTPRASPTGPTT 551
Oy 77 -----RLHEVCSVFPYDECTCYGSPFAGSAASTYPTENLLKKRGD 122
      |::|||::|||
Db 552 PPSPGGGVGAGAAWSRLINSFPGSPFRHRRMQVPTAEWSSLTBSSPELAKS- 610
Oy 121 LVEFVQAQYPMHAAVVGNGFQVNHRLREVINSFLTDAQGRGRGVNDLYRYKLSSA 180
      |::|||::|||
Db 611 -----WFGNF--ISLDKEQFLVLKD-----KPLSS-- 635
Oy 181 VFNALAHVAGKERELSMR--NSEFPAMCRGKREFFIGELRIQKOPRYLOQLSAPR 235
      |::|||::|||
Db 636 -IKADIVHAFSLPSLHSHVLSQTSFRA-----EYASGSPVFOQPVAFQVDISSE 687
Oy 239 S-----HTLEFQSLDLDIMEKRRNDQIGRAVLQELATHLHPAREEGD- 282
      |::|||::|||
Db 688 GPEPSPRRDGGGGGYSTVFTLLISG---PSRRFRVAVETIQAQDLSTHDPQVALADE 744
Oy 283 SNVART-----TPPRGR-PAFSSSEBDG 305
      |::|||::|||
Db 745 KNGAQTRPAGAPRRSLQPPRGRPDELSSSPRG 778

```

```

RESULT 7
US-10-612-783-4430
; Sequence 4430, Application US/10612783
; Publication No. US20060236419A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yinhua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53373) A
; CURRENT APPLICATION NUMBER: US/10/612,783
; CURRENT FILING DATE: 2003-07-02
; NUMBER OF SEQ ID NOS: 7098
; SEQ ID NO 4430
; LENGTH: 872
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(872)

```

```

OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4577_125772C.1.pap
US-10-612-783-4430

Query Match      6.9%; Score 115.5; DB 6; Length 872;
Best Local Similarity 21.1%; Pred. No. 0.41;
Matches 85; Conservative 54; Mismatches 123; Indels 141; Gaps 20;

QY 3 NOVEKLHLISYK-----VPTADPTGVDRDDGPRIGSVYIFSNDEDEVEPQPQPGDG 57
DB 429 HEIDEDSHRSHMEBEGHYEKDMQDEVDVADBDME-----SDNNRELKPEKXKVGPIG 483
QY 58 GLPDGDDGPPPPQ-----PPYDPRLHEVCSVFPYRDECTYKSPFAPSAA 103
DB 484 SAAPFGRRPARPRRGLVCGPDLGVDPKPFDPKTY-VEEDVFVTEGSGTKRIRLEDNI 542
QY 104 LSTTTPENL--LNKCKRGDLVEFVSQOYPRH-----AVVGNQOVVHLHRLVINSFLT 156
DB 543 VRMRTINNADGTTSC-----SNARFVKMDGTNQLIGN-----EVLDISVH 585
QY 157 DASQGR-----RGRVNDLYRYKPLSSAVVNALAHV----- 189
DB 586 EAHHDQSHLFLRHGKGLQSGRLHKM-RFMPSSLSKSHRLTLALVDSQNKTVGMOK 644
QY 190 -----GAKERELSWRNSESFAMCKRYGKREFKIGELRIGQPYRLQLSAQR 238
DB 645 WFTKDEKAMKEREYIG-QNIRAHSLQR--KRE-KVN--RKYTOPAPRRQLSP-- 695
QY 239 SHLEFQSLDLME-----KRNQOIGRAVLOGLATHLHPAEDEGDSN 284
DB 696 -----GFLDADLDEDEPGYDSRRMPGRRRFEDELENALEARRIINAK-----KSN 742
QY 285 VARTTP-----PPGRPP-----APSSSEEDGEAVAH 310
DB 743 MSRRVPHKPPYPARPARRQADEVSGSERSESEYETGEDIEH 785

RESULT 8
US-10-131-833A-364
; Sequence 364, Application US/10131833A
; Publication No. US20060257971A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gettisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACTIS ENCODING THE SAME
FILE REFERENCE: P3330R1C142
CURRENT FILING DATE: 2002-10-17
PRIOR APPLICATION NUMBER: US/40/131.833A
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17

```

```

PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 364
LENGTH: 328
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-833A-364

Query Match      6.1%; Score 101.5; DB 6; Length 328;
Best Local Similarity 28.5%; Pred. No. 1.7;
Matches 73; Conservative 24; Mismatches 72; Indels 87; Gaps 19;

QY 59 LPQGDGPPPPQPPY-DPRLHEVCSVFPYRDECTYKSPFAPSALSTYPPENLNKCK 117
DB 14 LRHQAQKPSPDAGPHQGRVHQA-----APLSDA----- 43
QY 118 PGDLVEFVSQOYPRHMAVYVGNFQVHLHRL--EVINSF--LT-DASQGRGRVNDLYR 172
DB 44 -----PHDDAH-GNFQYDHEAFLGREVAKEFDOLTPEESQARLGRIVDRMDR 89
QY 173 Y--KPLSSAVVNALAHGAKERELSWRNSESFAMCKRYG-KREFKIG-GEIRIGQP 227
DB 90 AGDDGCVSLAEIRAWIAH--TQQRH--RDSVS-AAADTYDDDRDGVGEBLR----- 139
QY 228 YRLQQLSAQRSHTL--EFQSLD-----LIMEKR--NDQIGRA-AVLQELATHL 273
DB 140 -----NATYGHYAPBEEFHVDVDAETKYKMLARDERRRVADQDDSDMATRELTAF 192
QY 274 HPAE-PEEGDSNVART 288
DB 193 HPEEPPHMRDIVIAET 208

RESULT 9
US-10-449-902-41174
; Sequence 41174, Application US/10449902
; Publication No. US20060123505A1
GENERAL INFORMATION:
APPLICANT: National Institute of Agricultural Sciences.
APPLICANT: Bio-Oriented Technology Research Advancement Institution.
APPLICANT: The Institute of Physical and Chemical Research.
APPLICANT: Foundation for Advancement of International Science.
FILE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
FILE REFERENCE: MOA-A0205Y1-US
CURRENT APPLICATION NUMBER: US/10/449.902
CURRENT FILING DATE: 2003-05-29
PRIOR APPLICATION NUMBER: JP 2002-203269
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: JP 2002-383870
PRIOR FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 56791
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 41174
LENGTH: 854
TYPE: PRT
ORGANISM: Oryza sativa
US-10-449-902-41174

Query Match      5.9%; Score 97.5; DB 6; Length 854;
Best Local Similarity 20.5%; Pred. No. 11;
Matches 70; Conservative 37; Mismatches 127; Indels 107; Gaps 16;

```

```
OY 13 YKEVPTADPTGVDRDDGPRIGSVYIFSNDDDEVEPQPPGDPGGGLPDC-GDGPPEPPOP 71
DB 151 YRPNPTIN-TSVQSMQGGPSGL-----PAQSSLPFGVGNHPPAASP 190
OY 72 -----QPYDRLHEVECSVFYRDECIYOKSFAPG-SAALSTYPPENLNKCKPGDLVE 123
DB 191 ISGPIVSABY-----SSNAAGASTSNAPFRPRPRSTSTSGEDTH 230
OY 124 FVS-----QAOYPRMAVYVGNFQVHLHRLVINSFLTDASQGRGRVNVNDLYR 172
DB 221 TVSSVTASQHVAVMAQSHSSIGPOQHMGISPTLNQSPITLHGAPGES----- 281
OY 173 YKPLSSAVRNALAHVAKERELSMRNSESFAAMRYCKREFK--IGSELATIGQPYRL 230
DB 282 ----TPSLTQSQFVH-GTRDRERSRDSATPG-----GRNTFKSVYGFVNSMDSVESA 331
OY 221 Q--IOLSAQ-----RSHTLFPOSL-----EDLIMEK--RRNDQIGRAVLOELAT 271
DB 332 OKKITISTPYDPVHLTHVGFNSDTGFTGLPKEMOOLLQESGISRDQANQAVMDIYA 391
OY 272 HLHPAPEEGDSNVAR-----TTPPGGRPPAPSSSEEDG 305
DB 392 FYQDATKSDGPTDVMKMGYAKGNNOAPGTPTDTSSSDG 432
```

RESULT 10

US-10-953-349-21087
Sequence 21087, Application US/10953349
Publication No. US20060107345A1

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
TITLE OF INVENTION: ENCODED THERBY
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 21087
LENGTH: 543
TYPE: PRT
ORGANISM: Glycine max
S-10-953-349-21087

Query Match 5.7%; Score 94.5; DB 6; Length 543;

Best Local Similarity 20.4%; Pred. No. 11;
Matches 78; Conservative 41; Mismatches 127; Indels 137; Gaps 16;

```
OY 18 TADPTGVDRDDGPRIGSVYIFSNDDDEVEPQ-----PQGPDPGGGLPDGGDGP--- 67
DB 80 TVATQGMARPETPGIS-SYKMGKTPEVTSSPVAVAGTSPPSAQBGFSGSNVARRPGQY 138
OY 68 PQGP-----QPYDRLHEVECSVFYRDECIYOKSFAP-----GSAALS 105
DB 139 PASPIKSSDQLVKDSKRVASVNGSDSDFGGDLFSASSFQPKQASSPQGRSSGTSALS 198
OY 106 -----TYPPENLNKCKPGDLVEFVSQAQYPRHAYVYVGNFQVHLHRLV 150
DB 199 SAIVPVGNGQSTRTSTPDSLQSLATQPVGAQLQQAQ-----PVVKQDQHAS 247
OY 151 INSFLTDSQGRGRVNDLYRKPPLSSAVRNALAHVAKERELSMRNSESFAAMWC- 209
DB 248 VQTHNKNPNSGLPGR-----LQDSA-----SSQSQAPWPRM 278
OY 210 -----YGK-----REFKIGSE-----LRIGQPYRLQIQLSAQSHLTLEFOS 246
DB 279 TQTDVQKWKVMEVDTDRDGKITGEQARNLFLSMRLPREVLKQVMDLSDQNDMSH--LS 336
OY 247 LEDL-----IMEKRNDQIGRAA-----VLOELATHLHPA-----EPEEGDS 283
DB 337 LREFCIALYMERHREGRLPAVLPSNIVLDLPTTGQPAAHYSSMGNPSAFOQOPETGTS 396
```

```
OY 284 NVARTTPPPRRPAPSSSEEDGE 306
DB 397 GARQVNPAGRPBRPRAAVSQSDE 419
```

RESULT 11

US-10-953-349-21086
Sequence 21086, Application US/10953349
Publication No. US20060107345A1

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai et al.
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDE
TITLE OF INVENTION: ENCODED THERBY
FILE REFERENCE: 2750-1579PUS2
CURRENT APPLICATION NUMBER: US/10/953,349
CURRENT FILING DATE: 2004-09-30
NUMBER OF SEQ ID NOS: 40252
SOFTWARE: PatentIn version 3.3
SEQ ID NO 21086
LENGTH: 586
TYPE: PRT
ORGANISM: Glycine max
US-10-953-349-21086

Query Match 5.7%; Score 94.5; DB 6; Length 586;

Best Local Similarity 20.4%; Pred. No. 12;
Matches 78; Conservative 41; Mismatches 127; Indels 137; Gaps 16;

```
OY 18 TADPTGVDRDDGPRIGSVYIFSNDDDEVEPQ-----PQGPDPGGGLPDGGDGP--- 67
DB 123 TVATQGMARPETPGIS-SYKMGKTPEVTSSPVAVAGTSPPSAQBGFSGSNVARRPGQY 181
OY 68 PQGP-----QPYDRLHEVECSVFYRDECIYOKSFAP-----GSAALS 105
DB 182 PASPIKSSDQLVKDSKRVASVNGSDSDFGGDLFSASSFQPKQASSPQGRSSGTSALS 241
OY 106 -----TYPPENLNKCKPGDLVEFVSQAQYPRHAYVYVGNFQVHLHRLV 150
DB 242 SAIVPVGNGQSTRTSTPDSLQSLATQPVGAQLQQAQ-----PVVKQDQHAS 290
OY 151 INSFLTDSQGRGRVNDLYRKPPLSSAVRNALAHVAKERELSMRNSESFAAMWC- 209
DB 291 VQTHNKNPNSGLPGR-----LQDSA-----SSQSQAPWPRM 321
OY 210 -----YGK-----REFKIGSE-----LRIGQPYRLQIQLSAQSHLTLEFOS 246
DB 322 TQTDVQKWKVMEVDTDRDGKITGEQARNLFLSMRLPREVLKQVMDLSDQNDMSH--LS 379
OY 247 LEDL-----IMEKRNDQIGRAA-----VLOELATHLHPA-----EPEEGDS 283
DB 380 LREFCIALYMERHREGRLPAVLPSNIVLDLPTTGQPAAHYSSMGNPSAFOQOPETGTS 439
OY 284 NVARTTPPPRRPAPSSSEEDGE 306
DB 440 GARQVNPAGRPBRPRAAVSQSDE 462
```

RESULT 12

US-11-056-355B-81762
Sequence 81762, Application US/11056355B
Publication No. US20060150283A1

GENERAL INFORMATION:

APPLICANT: Brovet, Vyacheslav
APPLICANT: Alexandrov, Nikolai
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 2750-1590PUS2
CURRENT APPLICATION NUMBER: US/11/056,355B
CURRENT FILING DATE: 2005-02-14
PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO 81762

LENGTH: 1453
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(1453)
OTHER INFORMATION: Ceres Seq. ID no. 12661234
US-11-056-355B-81762

Query Match 5.7%: Score 94.5; DB 7; Length 1453;
Best Local Similarity 23.2%; Pred. No. 37;
Matches 59; Conservative 25; Mismatches 83; Indels 87; Gaps 14;

QY 3 NOVEK-----LTHSYKEVPTADPTGVDRDGPRIQVSYFNSDDEVEPOPP-----Q 52
DB 1108 NSVEKGSRLILQOQKEVLPALPS-----GFLGDFVLMKGSVN---NPPAAMSS 1158
QY 53 GPDGCGLPDG-----GDGPPPOPOP-----YDRLHEVCSVYRDECIYOK 95
DB 1159 GPKKSTKPSRLRDIKEQEKMTTSSHPSPVPTQKAIPQAHQGS-----WSR 1210
Y 96 SFAPGSAALSTYPTENLNCKCPGD-----LVEFVSGAOYPHNAVYVGNFQVHLH 146
Y 1211 SASPSQAVSQSSSQ---SKSGDDDLFWGFEVOSTDPTKQGDFFH----- 1253
Y 147 RLEVINSLTDASQGRGRVNV---DLRYKPLS-----SSAVRNALAHGAKERE 195
DB 1254 -LTSQNSWGTNKTNGK---VNAGTSLNRQKSVMSGADRYLSSPVVTQA-SHGKKEAV 1307
QY 196 LSWNSESFAAWCR 209
DB 1308 TKLTEANGFRDWCCK 1321

RESULT 13

US-11-056-355B-81761
Sequence 81761, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT FILING DATE: 2005-02-14
PRIORITY FILING DATE: 2004-02-13
PRIORITY FILING DATE: 2004-02-13
NUMBER OF SEQ. ID NOS: 119966
SEQ. ID NO 81761
LENGTH: 1661
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(1661)
OTHER INFORMATION: Ceres Seq. ID no. 12661233
US-11-056-355B-81761

Query Match 5.7%: Score 94.5; DB 7; Length 1661;
Best Local Similarity 23.2%; Pred. No. 43;
Matches 59; Conservative 25; Mismatches 83; Indels 87; Gaps 14;

QY 3 NOVEK-----LTHSYKEVPTADPTGVDRDGPRIQVSYFNSDDEVEPOPP-----Q 52
DB 1316 NSVEKGSRLILQOQKEVLPALPS-----GFLGDFVLMKGSVN---NPPAAMSS 1366
QY 53 GPDGCGLPDG-----GDGPPPOPOP-----YDRLHEVCSVYRDECIYOK 95
DB 1367 GPKKSTKPSRLRDIKEQEKMTTSSHPSPVPTQKAIPQAHQGS-----WSR 1418
QY 96 SFAPGSAALSTYPTENLNCKCPGD-----LVEFVSGAOYPHNAVYVGNFQVHLH 146

DB 1419 SASPSQAVSQSSSQ---SKSGDDDLFWGFEVOSTDPTKQGDFFH----- 1461
QY 147 RLEVINSLTDASQGRGRVNV---DLRYKPLS-----SSAVRNALAHGAKERE 195
DB 1462 -LTSQNSWGTNKTNGK---VNAGTSLNRQKSVMSGADRYLSSPVVTQA-SHGKKEAV 1515
QY 196 LSWNSESFAAWCR 209
DB 1516 TKLTEANGFRDWCCK 1529

RESULT 14

US-11-056-355B-81760
Sequence 81760, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT FILING DATE: 2005-02-14
PRIORITY FILING DATE: 2004-02-13
PRIORITY FILING DATE: 2004-02-13
NUMBER OF SEQ. ID NOS: 119966
SEQ. ID NO 81760
LENGTH: 1714
TYPE: prt
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(1714)
OTHER INFORMATION: Ceres Seq. ID no. 12661232
US-11-056-355B-81760

Query Match 5.7%: Score 94.5; DB 7; Length 1714;
Best Local Similarity 23.2%; Pred. No. 45;
Matches 59; Conservative 25; Mismatches 83; Indels 87; Gaps 14;

QY 3 NOVEK-----LTHSYKEVPTADPTGVDRDGPRIQVSYFNSDDEVEPOPP-----Q 52
DB 1369 NSVEKGSRLILQOQKEVLPALPS-----GFLGDFVLMKGSVN---NPPAAMSS 1419
QY 53 GPDGCGLPDG-----GDGPPPOPOP-----YDRLHEVCSVYRDECIYOK 95
DB 1420 GPKKSTKPSRLRDIKEQEKMTTSSHPSPVPTQKAIPQAHQGS-----WSR 1471
QY 96 SFAPGSAALSTYPTENLNCKCPGD-----LVEFVSGAOYPHNAVYVGNFQVHLH 146
DB 1472 SASPSQAVSQSSSQ---SKSGDDDLFWGFEVOSTDPTKQGDFFH----- 1514
QY 147 RLEVINSLTDASQGRGRVNV---DLRYKPLS-----SSAVRNALAHGAKERE 195
DB 1515 -LTSQNSWGTNKTNGK---VNAGTSLNRQKSVMSGADRYLSSPVVTQA-SHGKKEAV 1568
QY 196 LSWNSESFAAWCR 209
DB 1569 TKLTEANGFRDWCCK 1582

RESULT 15

US-11-056-355B-1197
Sequence 1197, Application US/11056355B
Publication No. US20060150283A1
GENERAL INFORMATION:
APPLICANT: Brover, Vyacheslav
TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
FILE REFERENCE: 2750-1590PUS2
CURRENT FILING DATE: 2005-02-14

PRIOR APPLICATION NUMBER: 60/544,190
PRIOR FILING DATE: 2004-02-13
NUMBER OF SEQ ID NOS: 119966
SEQ ID NO: 1197
LENGTH: 219
TYPE: prt
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: peptide
LOCATION: (1)..(219)
OTHER INFORMATION: Ceres Seq. ID no. 12350568
US-11-056-355B-1197

Query Match 5.6%; Score 93.5; DB 7; Length 219;
Best Local Similarity 24.2%; Pred. No. 4.4;
Matches 45; Conservative 14; Mismatches 68; Indels 59; Gaps 7;

Qy	161	GRRGVVNDLYRYKPLSSAVVNNALA-----HVAKERELSWRNSESFAWCR	209
Db	39	GRRRVAVCCCGWERAGGAAAGAAARCGLVGCRHAGGPRGR--RRRRBAGWRR	95
Qy	210	YGRREFKIGELRIG---KQPYRLQIQLSAQSHITLFEQSLDLMKRRNDQIGRAV	265
Db	96	QGRRAAAVQEPBPGGAEEAEOAALQAVARPEHLQ-----DGRGDBPRGRRL	146
Qy	266	LQELATLHPAPEPEGDSNVA-----RTTPPG-----RPPAPS	299
Db	147	-----HKGPAEPGCAAGACGPGGRRRRPRPPGPPAAGPGGAGERRVAHAHEPPAPA	200
Qy	300	SEEDG 305	
Db	201	RRDEG 206	

Search completed: December 6, 2006, 21:55:07
Job time : 41 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 21:41:06; Search time 303 Seconds

(without alignments)
946.386 Million cell updates/sec

Title: US-10-736-227-1

Percent score: 1663
Sequence: 1 MGNOVEKLTLSYKEVPTAD.....PPGRPPAPSSSEEDGEAVAH 310

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: UniProt 7.2.*
1: uniprot_sprot.*
2: uniprot_tr embl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1663	100.0	310	096KNI_HUMAN	096KNI homo sapien
2	916	55.1	260	04RFD0_TERNING	04RFD0 tetraodon n
3	859.5	51.7	270	0568P3_BRARE	0568P3 brachydanio
4	642.5	38.6	292	086U22_HUMAN	086U22 homo sapien
5	642.5	38.6	292	087AM5_HUMAN	087AM5 homo sapien
6	642.5	38.6	292	096KNI_HUMAN	096KNI homo sapien
7	637.5	38.3	292	096E50_MOUSE	096E50 mus musculu
8	636.5	38.3	292	099K34_MOUSE	099K34 mus musculu
9	625	37.6	297	032CAL_BOVIN	032CAL bos taurus
10	599	36.0	260	06PGW5_BRARE	06PGW5 brachydanio
11	581.5	35.0	304	04S116_TERNING	04S116 tetraodon n
12	269.5	16.2	354	08NAB7_HUMAN	08NAB7 homo sapien
13	250.5	15.1	201	071RD3_HUMAN	071RD3 homo sapien
14	136	8.2	218	04RPF7_TERNING	04RPF7 tetraodon n
15	136	8.2	164	TIG3_HUMAN	09UL19 homo sapien
16	133.5	8.0	235	08T778_BRARE	08T778 brachydanio
17	131	7.9	228	06P884_XENTR	06P884 xenopus tro
18	129.5	7.8	228	06INW8_XENTR	06INW8 xenopus lae
19	122.5	7.4	158	06PA17_XENTR	06PA17 xenopus lae
20	122	7.3	732	04UBD4_THERAN	04UBD4 theileria a
21	122	7.3	794	BRSL1_HUMAN	08UDC3 homo sapien
22	120	7.2	1219	090710_CHICK	090710 gallus gall
23	118	7.1	162	HRSL3_MOUSE	08XJUI mus musculu
24	118	7.1	162	03V3C3_MOUSE	03V3C3 m 3 days ne
25	117	7.0	162	HRSL2_HUMAN	09NWW9 homo sapien
26	117	7.0	162	HRSL3_HUMAN	053116 homo sapien
27	117	7.0	162	HRSL3_MOUSE	035Y13 pongo pygma
28	117	7.0	162	035Y13_HUMAN	035Y13 homo sapien
29	117	7.0	213	05N1P9_SYNP6	05N1P9 synchococ
30	116.5	7.0	167	HRSL1_MOUSE	09GQU4 mus musculu
31	116.5	7.0	167	Q5D099_MOUSE	Q5D099 mus musculu

ALIGNMENTS

```

RESULT 1
ID 096KNI_HUMAN PRELIMINARY; PRT; 310 AA.
AC 096KNI;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DE 07-FEB-2006, entry version 15.
DE NSE2 protein (Breast cancer membrane protein 101).
GN Name=NSE2; Synonyms=PAM84B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo
OX NCBI_TaxId=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hughes P.J., Stanway G.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltan E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.T., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC
DR EMBL; AJ417849; CAD10629.1; -; mRNA.
DR EMBL; BC052957; AAHS2957.1; -; mRNA.
DR Ensemble; ENSG00000168672; Homo sapiens.

```

DR GO: GO:0005737; C:cytoplasm; IDA.
 DR GO: GO:0005886; C:plasma membrane; IDA.
 DR GO: GO:0005515; F:protein binding; IPI.
 SQ SEQUENCE 310 AA; 34474 MW; B587A3F9CB41938 CRC64;

Query Match 100.0%; Score 1663; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1,1e-119; Indels 0; Gaps 0;
 Matches 310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGNQVEKTLTHSYKEVPPTADPTGVDRDDGPRIGVSYFISNDDEDEVEPPQPPGPGGGGLP 60
 DB 1 MGNQVEKTLTHSYKEVPPTADPTGVDRDDGPRIGVSYFISNDDEDEVEPPQPPGPGGGGLP 60
 QY 61 DGGGPPPPQPPQPPDPRRLHEVECSFYRDECTIQKSPAPGSAALSTYPENLNKCKRGD 120
 DB 61 DGGGPPPPQPPQPPDPRRLHEVECSFYRDECTIQKSPAPGSAALSTYPENLNKCKRGD 120
 QY 121 LVEFVSQAQYPHMAVYVGNFQVHLHRLLEVINSFLTDASQGRGRVNDLYRYKPLSSSA 180
 DB 121 LVEFVSQAQYPHMAVYVGNFQVHLHRLLEVINSFLTDASQGRGRVNDLYRYKPLSSSA 180
 QY 181 VVRNALAHVGAKEKRELSPWNSSESPFAWCRYGKREPKIGELRIKQPYRLQIQLSAQRSH 240
 DB 181 VVRNALAHVGAKEKRELSPWNSSESPFAWCRYGKREPKIGELRIKQPYRLQIQLSAQRSH 240
 QY 241 TLEFQSLLEDLMKRRNDQIGRAAVLOELATHLHRAPEEGDSNVAKTTTPPGRPAPASS 300
 DB 241 TLEFQSLLEDLMKRRNDQIGRAAVLOELATHLHRAPEEGDSNVAKTTTPPGRPAPASS 300
 QY 301 EEDGEAVAH 310
 DB 301 EEDGEAVAH 310

RESULT 2
 Q4RFD0_TETNG PRELIMINARY; PRT; 260 AA.
 ID Q4RFD0_TETNG PRELIMINARY; PRT; 260 AA.
 AC Q4RFD0_TETNG PRELIMINARY; PRT; 260 AA.
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 07-FEB-2006, entry version 3.
 DE Chromosome 8 SCAPI5119, whole genome shotgun sequence. (Fragment).
 GN ORFNames=CGTNG00035376001;
 OS Tetradodon nigroviridis (Green puffer).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 CC Tetraodontidae; Tetraodontidae; Tetraodon.
 CX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15496914; DOI=10.1038/nature03025;
 RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 RA Maucel E., Bouneau L., Fischer C., Ozouf-Costez C., Bernot A.,
 RA Nicod S., Jaffe D., Fisher S., Lutfalla N., Dossat C., Segurens B.,
 RA Desliva C., Salenouat M., Levy M., Boudet N., Castellan S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Blemont C., Skalli Z., Cateolico L., Poulain J., De Bernardis V.,
 RA Crnaud C., Duprat S., Broctier P., Coutanceau J.-P., Gouy J.,
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lauder V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
 RA "Genome duplication in the teleost fish Tetradodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope, Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 DR EMBL: CAE01015119; CAG12902.1; -; Genomic DNA.
 FT NON TER 260
 SQ SEQUENCE 260 AA; 29993 MW; 5B93F74B89EC5E34 CRC64;

Query Match 55.1%; Score 916; DB 2; Length 260;
 Best Local Similarity 62.5%; Pred. No. 2.9e-62;
 Matches 172; Conservative 42; Mismatches 45; Indels 16; Gaps 4;

QY 1 MGNQVEKTLTHSYKEVPPTADPTGVDR-DDGRIGVSYFISNDDEDEVEPPQPPGPGGGGL 59
 DB 1 MGNQVEKTLTHSYKEVPPTADPTGVDR-DDGRIGVSYFISNDDEDEVEPPQPPGPGGGGL 51
 QY 60 PDGDDGPPPPQPPQPPDPRRLHEVECSFYRDECTIQKSPAPGSAALSTYPENLNKCKRG 119
 DB 52 LDHFOAMSNHEKRFDDQ-DELECAHYRECYER-----RSAATHSAESLNKCKRG 105
 QY 120 DLVEFVSQAQYPHMAVYVGNFQVHLHRLLEVINSFLTDASQGRGRVNDLYRYKPLSSS 179
 DB 106 DLLEFVATGQYPHMAVYVGNFQVHLHRLLEVINSFLTDASQGRGRVNDLYRYKPLSSS 165
 QY 180 AVVRNALAHVGAKEKRELSPWNSSESPFAWCRYGKREPKIGELRIKQPYRLQIQLSAQRSH 239
 DB 166 VVRNALAHVGAKEKRELSPWNSSESPFAWCRYGKREPKIGELRIKQPYRLQIQLSAQRSH 225
 QY 240 HTLEFQSLLEDLMKRRNDQIGRAAVLOELATHLH 274
 DB 226 HLEFQSLLEDLMKRRNDQIGRAAVLOELATHLH 260

RESULT 3
 Q56BP3_BRARE PRELIMINARY; PRT; 270 AA.
 ID Q56BP3_BRARE PRELIMINARY; PRT; 270 AA.
 AC Q56BP3_BRARE PRELIMINARY; PRT; 270 AA.
 DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
 DT 10-MAY-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Hypothetical protein zgc:110182.
 GN ORFNames=zgc:110182;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinodontiformes;
 CC Cyprinidae; Danio.
 CX NCBI_TaxID=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Mixed tissue.
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Falingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Ditschenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M., Udén T.B., Tothilyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bock S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettlemann W., Nadeau A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mixed tissue.
RG NIH MGC Project;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
CC EMBL: BC092778; AAH92778.1; -, mRNA.
CC ZFIN: ZDB-GENE-050417-381; zgc:110182.
CC Hypothetical protein.
SQ SEQUENCE 270 AA; 31111 MW; 124E589083BCD936 CRC64;

Query Match 51.7%; Score 859.5; DB 2; Length 270;
Best Local Similarity 59.4%; Pred. No. 6,9e-58;
Matches 168; Conservative 44; Mismatches 52; Indels 19; Gaps 5;

QY 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSVYSFSDNDEVDVPPPPQGGGGL 59
DB 1 MGNQVEKLTHTLNNELPTADPNGFDEPDAPRIGSVYISADDEG----- 47
QY 60 PDGCGPPPPQPPDPRLEHVECSYFDECTYQKSPAGSAALSTYPPENLKKCKPG 119
DB 48 -ENMDETEDVAVHDCR-NELECAVYRDEVYERNSRFG--VGLSLRENLANCKAG 103
QY 120 DVEPFSQAYPHMAVYVGNFQVHLRLLEVINSLPTDASQGRGRGVNDLYRKLSS 179
DB 104 DLVEFALIQPFPMACVGELOVHLRNEIKCDFLDASQGRGRIVNELYKFRLLSD 163
QY 180 AVVRNALAHVAGKERELSWRNSESFAAMCRYGKREKIGELRIGQPRLOLSAORS 239
DB 164 AVQNMMEQVGLKERVCKKSECFAMCRFGKREKIGELRIGQPRKMLQLSEKKS 223
QY 240 HTLEFQSLDLIMEKRRNDQIGRAVLQELATLHLPAPPEEGD 282
DB 224 HTLEFQSLDLIMEKRRNDQIGRAVLQELATLHLPAPPEEGD 265

RESULT 4
Q86U22 HUMAN PRELIMINARY; PRT; 292 AA.
ID Q86U22;
AC Q86U22;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 12.
DE Hypothetical protein NSE1.
GN Name=NSE1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22388297; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins B., Bennett K.H., Schenker C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheet C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Umed T.B., Toshiki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.N., Krzywicki M.I., Skalska U., Smolius D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN (3)
RP NUCLEOTIDE SEQUENCE.
RA Spalding L., Haakenson W., Swearingen S.;
RT "The sequence of Homo sapiens BAC clone RP11-314E10.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN (4)
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN (5)
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN (6)
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
CC EMBL: BC052284; AAH52284.1; -, mRNA.
CC EMBL: AC011897; AAY15041.1; -, Genomic DNA.
CC EMBL: ENSG00000162981; Homo sapiens.
CC Hypothetical protein.
SQ SEQUENCE 292 AA; 32491 MW; 652F11029F517626 CRC64;

Query Match 38.6%; Score 642.5; DB 2; Length 292;
Best Local Similarity 46.4%; Pred. No. 3,9e-41;
Matches 137; Conservative 36; Mismatches 87; Indels 35; Gaps 7;

QY 1 MGNQVEKLTHTLSYKEVPTADPTGVDRDDGPRIGSVYSFSDNDEVDVPPPPQGGGGL 60
DB 1 MGNQDRITHTLNNELPTADPNFSGIEKDE-LRVGVAIFSDDEDL-----DERGQ 50
QY 61 D--GGDPPPPQPPDPRLEH-----EVECSYFVDECTYQK-SFAPGSAALST 106
DB 51 DKRGVAPPGCTPCPSPSRHHHLLHQLVLMETQSSARGCICFSKXSGCGQADLAV 110
QY 107 YTPENLKKCKPQDLVEFY-----SQAYPHMAVYVGNFQVHLRLLEVINSLPTDAS 159
DB 111 YAVTALPACERQDLLELMLQPAPEPAPAPAHMAVYVGGQIILHGCETIRQDSLYEAG 170
QY 160 QGRGRGVNDLYRKLSSAVVRNALAHVAGKERELSWRNSESFAAMCRYGKREKIG 219
DB 171 AAVNGRVNWSYRPLVAVELVQVNAAGHLGKSEETCTNSESFAAMCRFGKREKIG 230
QY 220 ELRIGKQPP-----YRLDQLSAORSHTLEFQSLDLIMEKRRNDQIGRAVLQELA 270
DB 231 EYPAGTQPPQOQYLYKLVHGENKVTAFRPHSLEDLIREKRIIDASGRVLQELA 285

RESULT 5
Q8TAM5 HUMAN PRELIMINARY; PRT; 292 AA.
ID Q8TAM5;
AC Q8TAM5;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 13.
DE FAM84A protein.
GN Name=FAM84A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Hunkeler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., USCIN T.B., Toshiyuki S., Carinci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RG NIH MGC Project;
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL: BC026346; AAH26346.1; -; mRNA.
 DR Ensembl; ENSG00000162981; Homo sapiens.
 DR SEQUENCE 292 AA; 32532 MW; EAJF807C1F469F56 CRC64;
 SQ
 Query Match 38.6%; Score 642.5; DB 2; Length 292;
 Best Local Similarity 46.4%; Pred. No. 3.9e-41;
 Matches 137; Conservative 36; Mismatches 87; Indels 35; Gaps 7;
 QY 1 MGNQVEKTLHSLYKVEPTADPTGVDPRIGVSYFNSNDEVEPQPPQGGGLP 60
 DB 1 MGNQIDRTHLNLVSELPTGDPGIEKDE-LRVGVAFFSDDEBDL-----DERQOP 50
 QY 61 D-GGDDPPPPQPPQPPQPPRLH-----EVECSYFYRDECIYOK-SFAPGSAALST 106
 DB 51 DKFGVKAPEGCTPCBPESRHHHLLHQLVLNETQFSAPRQECIFSKVSGGPGQADLSV 110
 QY 107 YTPENLNKCKRPGDLVEFV-----SOAYPHMAVYVGNFQVHLHLEVINSLFDAS 159
 DB 111 YAVTALPALCEPGLLELMTQPAPEPPAPAPMAVYVGGQIITHLHGEIRQDSLYEAG 170
 QY 160 QGRGRVYNDLYRKYPLSSSAVYRNALAHVAKERELSMRSESPAAACRYKREPKFG 219
 DB 171 AANVGRVYNSWYRRPLVAELVQNAACHLGLKSEELCWTNSESFAACRCRKGKPKXGG 230
 QY 220 ELRTGKOP-----YRLQIOLSAQRSHLTFOSLEDLIMEKRNDQIGRAVLOELA 270
 DB 231 EVPAQTQPPQOQYLYKVLHGENKVTATFHSLEDLIREKRNDASGRRLRVLOELA 285
 RESULT 6
 Q96KN4 HUMAN PRELIMINARY; PRT; 292 AA.
 AC Q96KN4
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 8.
 DE NSE1 protein.
 GN Name=NSE1;
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RA Hughes P.J., Stanway G.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL: AJ417080; CAD10038.1; -; mRNA.
 DR Ensembl; ENSG00000162981; Homo sapiens.
 DR SEQUENCE 292 AA; 32482 MW; BA55B6AD3FE345A CRC64;
 SQ
 Query Match 38.6%; Score 642.5; DB 2; Length 292;
 Best Local Similarity 46.4%; Pred. No. 3.9e-41;
 Matches 137; Conservative 36; Mismatches 87; Indels 35; Gaps 7;
 QY 1 MGNQVEKTLHSLYKVEPTADPTGVDPRIGVSYFNSNDEVEPQPPQGGGLP 60
 DB 1 MGNQIDRTHLNLVSELPTGDPGIEKDE-LRVGVAFFSDDEBDL-----DERQOP 50
 QY 61 D-GGDDPPPPQPPQPPQPPRLH-----EVECSYFYRDECIYOK-SFAPGSAALST 106
 DB 51 DKFGVKAPEGCTPCBPESRHHHLLHQLVLNETQFSAPRQECIFSKVSGGPGQADLSV 110
 QY 107 YTPENLNKCKRPGDLVEFV-----SOAYPHMAVYVGNFQVHLHLEVINSLFDAS 159
 DB 111 YAVTALPALCEPGLLELMTQPAPEPPAPAPMAVYVGGQIITHLHGEIRQDSLYEAG 170
 QY 160 QGRGRVYNDLYRKYPLSSSAVYRNALAHVAKERELSMRSESPAAACRYKREPKFG 219
 DB 171 AANVGRVYNSWYRRPLVAELVQNAACHLGLKSEELCWTNSESFAACRCRKGKPKXGG 230
 QY 220 ELRTGKOP-----YRLQIOLSAQRSHLTFOSLEDLIMEKRNDQIGRAVLOELA 270
 DB 231 EVPAQTQPPQOQYLYKVLHGENKVTATFHSLEDLIREKRNDASGRRLRVLOELA 285
 RESULT 7
 Q9D650 MOUSE PRELIMINARY; PRT; 292 AA.
 AC Q9D650
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 19.
 DE 10 days neonate skin cDNA, RIKEN full-length enriched library,
 DE clone:4731402F03 product:NSE1 PROTEIN homolog.
 GN Name=AW125753;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Methods Enzymol. 303:19-44(1999).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=Skin;
 RC PubMed=16141072; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Allen J.E., Allen J.E., Allen J.E., Bailey T.L.,
 RA Ambesi-Impombato A., Apweiler R., Auraliyya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,

RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dairymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georjic-Hemming P., Gigeras T.R., Gojobori T., Green R.E.,
 RA Guerinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummelbeck L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Keiso J., Kikuma H.,
 RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Motegut-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 RA Petrovsky N., Piazza S., Reid J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Sempé C.A., Seno S., Sessa L., Sheng Y.,
 RA Shilbata C., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
 RA Tammara K., Tan S.L., Tang S., Taylor M.S., Tegner J., Telchmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Bruscia V., Quackenbush J.,
 RA Wahlestedt C., Wattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arikawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawai H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimomiya N.,
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watanishi A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.,
 RA "The transcriptional landscape of the mammalian genome.",
 RA Science 309:1559-1563(2005).
 RA [3]
 RA NUCLEOTIDE SEQUENCE.
 RA STRAIN=C57BL/6J; TISSUE=Skin;
 RA PubMed=16141073; DOI=10.1126/science.1112009;
 RA RIKEN Genome Exploration Research Group, and Genome Science Group
 RA (Genome Network Core Team) and the PANTOM Consortium;
 RA "Antisense transcription in the mammalian Transcriptome.",
 RA Science 309:1564-1566(2005).
 RA [4]
 RA NUCLEOTIDE SEQUENCE.
 RA STRAIN=C57BL/6J; TISSUE=Skin;
 RA MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Oosato N., Saito R., Suzuki H., Yamana H., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirm L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani C., Frazer K.S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazee K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guerinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kanoy G.A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Masushima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempé C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynnaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasuniishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.",

RL Nature 420:563-573(2002).
 RN [5]
 RN NUCLEOTIDE SEQUENCE.
 RN STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana H.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirm L.M., Scandali P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Baren G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guerinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schonbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
 RA Wynnaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohetsuki S.,
 RA Hayashizaki Y.,
 RA "Functional annotation of a full-length mouse cDNA collection.",
 RT Nature 409:685-690(2001).
 RL [6]
 RL NUCLEOTIDE SEQUENCE.
 RL STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.",
 RL Genome Res. 10:1617-1630(2000).
 RL [7]
 RL NUCLEOTIDE SEQUENCE.
 RL STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kikunishi T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanishi M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-34-format
 RT sequencing pipeline with 384 multicapillary sequencer.",
 RL Genome Res. 10:1757-1771(2000).
 RL [8]
 RL NUCLEOTIDE SEQUENCE.
 RL STRAIN=C57BL/6J; TISSUE=Skin;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
 RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.,
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivative license
 CC EMBL: AK014625; BAB29470.1; -; mRNA.
 DR Ensembl: ENSMUSG0000020607; Mus musculus.
 DR MGI: MGI:2145011; AW125753.
 SQ SEQUENCE 292 AA; 32676 MW; D500C38DC50AFAFA CRC64;
 Query Match 38.3%; Score 637.5; DB 2; Length 292;

Best Local Similarity 56.9%; Pred. No. 26-36;
Matches 119; Conservative 21; Mismatches 42; Indels 27; Gaps 5;

QY 1 MGNQVEKTLTHLSYKEVPTADPTGVDRDDGPRIGVSYIFSNDEDEVEPOPPGPDGGGLP 60
Db 1 MGNQADLHLSLTAIEVPTVPCVDADGSRIGVSYIFSNDD-DEEPE----- 47
QY 61 DGGDPPPOPPQYDPLHVECSVFYRDECITYOKSPAFGSAALSTTPENLKNCKRPGD 120
Db 48 EGGASDSNOBEKLFDRDEDECAVFPYRDDCVYK--RARRASLEVYSPENLNCRAAGD 105
QY 121 LVEVSGAQQYPHAAVYVGNFQVHLHLEVINSLFTDASQGRGRGVVNDLYRYKPLSSA 180
Db 106 L-----YPHAAVYVGDFOVHLHRAEVKNSFLTDASGRGRGVVNDLYRYKPLSGEV 157
QY 181 VERNALAHVGAKEKELSMRNSSEFAMCR 209
Db 158 VVONAMEQVGLKODELSMRNSES-----CR 182

RESULT 12

Q8NAB7 HUMAN PRELIMINARY; PRT; 354 AA.

AC O8NAB7, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-MAR-2006, entry version 12.
DE CDNA FLJ35335 f18, clone PROSTR2015287.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Magatsuma M., Shikiori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Ariita M., Imose N.,
RA Mutsaers J. K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yochiya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamaguchi M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
RA Okawara A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Komatsu T.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T., Nakagawa K.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Yamashita K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi Y., Masuno Y., Yamashita K.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs";
RL Nat. Genet. 36:40-45(2004).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC EMBL; AK092654; BAC03938.1; -, mRNA.

DR Ensemble1; ENSG00000162981; Homo sapiens.
SQ SEQUENCE 354 AA; 37751 MW; DF510D1FEFF93CID CRC64;

Query Match 16.2%; Score 269.5; DB 2; Length 354;
Best Local Similarity 32.8%; Pred. No. 2.6e-12;
Matches 76; Conservative 25; Mismatches 80; Indels 51; Gaps 10;

QY 1 MGNQVEKTLTHLSYKEVPTADPTGVDRDDGPRIGVSYIFSNDEDEVEPOPPGPDGGGLP 60
Db 1 MGNQADLHLSLTAIEVPTVPCVDADGSRIGVSYIFSNDD-DEEPE----- 47
QY 61 DGGDPPPOPPQYDPLHVECSVFYRDECITYOKSPAFGSAALSTTPENLKNCKRPGD 120
Db 51 DKFGVKAAPGCTPCPESSPRHHHLHLQVLNFTQSFARQOECIFSVSGGPGQADLSV 110
QY 107 YTPENLKNCKRPGDLVEFV-----SQAQVPHAAVYVGNFQVHLHLEVINSLFTDAS 159
Db 111 YAVTALPALCEPGLLELMLQPAPEPPAPAPHAASRRR-SAGRTTRASPPGAAIASGS 169
QY 160 QGRGRGVVNDLYRYKPLSSAVVARNALAHVGAKEKELSMRNSSEFAMCR 209
Db 170 SRRGRC---RQASPRSSSTISR-----CTWERTTRSTPEGTAM 206

RESULT 13

Q71RD3 HUMAN PRELIMINARY; PRT; 201 AA.

AC Q71RD3, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Pp1517.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.Q., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC EMBL; AF370404; AA015240.1; -, mRNA.
SQ SEQUENCE 201 AA; 22213 MW; 74297F3D71EB39F CRC64;
Query Match 15.1%; Score 250.5; DB 2; Length 201;
Best Local Similarity 51.5%; Pred. No. 3.6e-11;
Matches 53; Conservative 10; Mismatches 35; Indels 5; Gaps 2;

QY 157 DASQGRGRGVVNDLYRYKPLSSAVVARNALAHVGAKEKELSMRNSSEFAMCRGKREFK 216
Db 75 EAGANVGRVYVNSWYRPLVAELVQNAACHLGLKSEIEICMTNSESPFAMCRFGKREFK 134
QY 217 IGGELRIKQYPRQIOLSA-----QRSHTLEFOSLEDIMEKR 255
Db 135 AGGEVPAFLAPAAVLSQAGARGQPPR-OFHSLDLIREKR 176

Oy	158	ASGRRGRVNDL-YRYKPLSSAVRNALAHGAKER-ELSMRNSEFPAACRYGKREF	215
Db	68	VVGCCYRVNNSLDHEYOPRPVEVVISSAKEMVGOIKMKYSIVSRNCEHFTQLRYGKGRG	127
Oy	216	K 216	
Db	128	K 128	

Search completed: December 6, 2006, 21:49:21
job time : 307 secs

```

; TITLE OF INVENTION: RETINOID INDUCED GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,303
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: ALRGN.062A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-808-303-12

Query Match      8.5%; Score 141; DB 1; Length 164;
Best Local Similarity 32.2%; Pred. No. 1.7e-06;
Matches 39; Conservative 21; Mismatches 39; Indels 22; Gaps 5;

QY 117 KPGDLVEFVSQAQYPHNAVYVGNFOVHL-----HRLVINSFLTDSQGRG---- 164
DB 9 KPGDLLE-IRLGEYHVALYIGGYVTHLAPPEYPCGAGSSVSFVLSNBAEVRGRLED 67

QY 165 -----RVVNDL-YRYKPLSSSAVVRNALAHVGAKER-ELSWRNSESFAWCRYGKREF 215
DB 68 VVGCCCVRVNNSLDHEQPRPVEVLIISAKEMVQKMKYSIVSNCEHFVAQLRYGKSRG 127

QY 216 K 216
DB 128 K 128

RESULT 3
US-08-996-533-12
; Sequence 12, Application US/08996533
; Patent No. 6294657
; GENERAL INFORMATION:
; APPLICANT: Nagpal, Sunil
; APPLICANT: Dissepio, Daniel
; APPLICANT: Chandraratna, Roshantha
; TITLE OF INVENTION: RETINOID INDUCED GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,533
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,303
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: ALRGN.062A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-996-533-12

Query Match      8.5%; Score 141; DB 2; Length 164;
Best Local Similarity 32.2%; Pred. No. 1.7e-06;
Matches 39; Conservative 21; Mismatches 39; Indels 22; Gaps 5;

QY 117 KPGDLVEFVSQAQYPHNAVYVGNFOVHL-----HRLVINSFLTDSQGRG---- 164
DB 9 KPGDLLE-IRLGEYHVALYIGGYVTHLAPPEYPCGAGSSVSFVLSNBAEVRGRLED 67

QY 165 -----RVVNDL-YRYKPLSSSAVVRNALAHVGAKER-ELSWRNSESFAWCRYGKREF 215
DB 68 VVGCCCVRVNNSLDHEQPRPVEVLIISAKEMVQKMKYSIVSNCEHFVAQLRYGKSRG 127

QY 216 K 216
DB 128 K 128

RESULT 4
US-08-808-303-8
; Sequence 8, Application US/08808303
; Patent No. 5776687
; GENERAL INFORMATION:
; APPLICANT: Nagpal, Sunil
; APPLICANT: Dissepio, Daniel
; APPLICANT: Chandraratna, Roshantha
; TITLE OF INVENTION: RETINOID INDUCED GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,303
; FILING DATE:
; CLASSIFICATION: 435
```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2006, 21:49:36 ; Search time 52 Seconds
(without alignments)
521.817 Million cell updates/sec

Title: US-10-736-227-1

Perfect score: 1663
Sequence: 1 MGNQVEKTLTHLSYKEVPTAD.....PCGRPPAPSSSEEDGEAVAH 310

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /EMC_Celerra_SIDS3/pcodata/2/1aa/5_COMB.pep.*
2: /EMC_Celerra_SIDS3/pcodata/2/1aa/6_COMB.pep.*
3: /EMC_Celerra_SIDS3/pcodata/2/1aa/7_COMB.pep.*
4: /EMC_Celerra_SIDS3/pcodata/2/1aa/H_COMB.pep.*
5: /EMC_Celerra_SIDS3/pcodata/2/1aa/PCUTUS_COMB.pep.*
6: /EMC_Celerra_SIDS3/pcodata/2/1aa/RB_COMB.pep.*
7: /EMC_Celerra_SIDS3/pcodata/2/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Hit No.	Score	Query Match	Length	DB ID	Description
1	269.5	16.2	354	2 US-10-104-047-3151	Sequence 3151, App1
2	141	8.5	164	1 US-08-808-303-12	Sequence 12, App1
3	141	8.5	164	2 US-08-996-533-12	Sequence 12, App1
4	136	8.2	164	1 US-08-808-303-8	Sequence 8, App1
5	136	8.2	164	2 US-08-996-533-8	Sequence 8, App1
6	136	8.2	164	2 US-08-801-742-1	Sequence 1, App1
7	122	7.3	703	2 US-10-116-326-6	Sequence 6, App1
8	122	7.3	703	2 US-10-803-377-6	Sequence 6, App1
9	122	7.3	762	2 US-10-116-326-4	Sequence 4, App1
10	122	7.3	762	2 US-10-803-277-4	Sequence 4, App1
11	122	7.3	778	2 US-10-116-326-2	Sequence 2, App1
12	122	7.3	778	2 US-10-003-690-2	Sequence 2, App1
13	122	7.3	778	2 US-10-803-277-2	Sequence 2, App1
14	116	7.0	160	2 US-08-801-742-4	Sequence 4, App1
15	114	6.9	162	2 US-08-801-742-3	Sequence 3, App1
16	113	6.8	522	2 US-10-104-047-3665	Sequence 3665, App1
17	109	6.6	603	2 US-09-930-181-17	Sequence 17, App1
18	109	6.6	614	2 US-10-283-247-9	Sequence 9, App1
19	109	6.6	614	2 US-10-283-247-10	Sequence 10, App1
20	109	6.6	636	2 US-10-283-247-5	Sequence 5, App1
21	109	6.6	674	2 US-10-283-247-2	Sequence 2, App1
22	109	6.6	674	2 US-10-283-247-7	Sequence 7, App1
23	109	6.6	674	2 US-10-283-247-8	Sequence 8, App1
24	103.5	6.2	572	2 US-09-543-681A-7137	Sequence 7137, App1
25	102.5	6.2	262	2 US-08-311-731A-333	Sequence 333, App1
26	101.5	6.1	328	1 US-08-828-242-1	Sequence 1, App1

27	101.5	6.1	328	1 US-08-910-927B-1	Sequence 1, App1
28	101.5	6.1	328	2 US-09-206-499-1	Sequence 1, App1
29	101.5	6.1	328	2 US-09-370-270-1	Sequence 1, App1
30	101.5	6.1	328	2 US-09-507-794A-221	Sequence 221, App1
31	101.5	6.1	328	2 US-09-505-125A-221	Sequence 221, App1
32	101.5	6.1	328	2 US-09-502-775A-221	Sequence 221, App1
33	101.5	6.1	328	2 US-09-506-700-221	Sequence 221, App1
34	101.5	6.1	328	2 US-09-503-603A-221	Sequence 221, App1
35	101.5	6.1	328	2 US-09-504-920A-221	Sequence 221, App1
36	101.5	6.1	328	2 US-09-509-064-221	Sequence 221, App1
37	101.5	6.1	328	2 US-09-505-381A-221	Sequence 221, App1
38	101.5	6.1	328	2 US-09-506-618-221	Sequence 221, App1
39	101.5	6.1	328	2 US-09-506-646-221	Sequence 221, App1
40	101.5	6.1	328	2 US-09-504-462-221	Sequence 221, App1
41	101.5	6.1	328	2 US-09-502-736A-221	Sequence 221, App1
42	101.5	6.1	328	2 US-09-506-722A-221	Sequence 221, App1
43	101.5	6.1	328	2 US-09-505-449-221	Sequence 221, App1
44	101.5	6.1	328	2 US-09-503-562B-221	Sequence 221, App1
45	101.5	6.1	328	2 US-09-506-679A-221	Sequence 221, App1

ALIGNMENTS

```

RESULT 1
US-10-104-047-3151
Sequence 3151, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104, 047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3151
LENGTH: 354
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-3151

Query Match 16.2%; Score 269.5; DB 2; Length 354;
Best Local Similarity 32.8%; Pred. No. 2.2e-19;
Matches 76; Conservative 25; Mismatches 80; Indels 51; Gaps 10;

QY 1 MGNQVEKTLTHLSYKEVPTADPTGVDRDDGPRIGSVYIFSNDDDEVPQPPGGGLP 60
DB 1 MGNQDRITHTLANSLEPTGDSGIEKDE-LRQVAVAFSSDEEDL-----DERGQF 50
QY 61 D--GGDPPPPPOPPYDPRLLH-----EVEGVPFRDCTIYK-SFAPGSALST 106
DB 51 DKFGVAPPCCTCPSPSRHHHLLHQLVLTNTPSFAFGQBCIFSKVGGPGGALSV 110
QY 107 YTEPNLNNCKPGDIVEFV-----SQAQYPMNAVVGNGFQVHLRLVINSFLTDA 159
DB 111 YAVTALPCLCEPDLELMLQAPAPPPAPAPMAASAR-SAGRTTRAPPPAALASGS 169
QY 160 QGRGRGVNDLVYRKKPLSSGAVVNNALAHGAKEREISWNSSES---FAAW 207
DB 170 SRREGRC---RQARSPRSSSTISR-----CTWERTSTPPGFTAW 206

RESULT 2
US-08-808-303-12
Sequence 12, Application US/08808303
Patent No. 5776687
GENERAL INFORMATION:
APPLICANT: Nagpal, Sunil
APPLICANT: Disepio, Daniel
APPLICANT: Chandiracna, Roshantha

```


Db 737 GL 738

RESULT 9
formin - mouse
S11515

C:Species: Mus musculus (house mouse)

C>Date: 22-Jan-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: S11515

R:Moynick, R.P.; as, R.L.; Zeller, R.; Vogt, T.F.; Leder, P.

Nature 346, 850, 1990

A:Title: 'Formin': proteins deduced from the alternative transcripts of the limb deform

A:Reference: 'Formin': S11515; MUID:90363291; PMID:2392150

A:Accession: S11515

A:Molecule: mRNA

A:Residues: 1-1468 <MO>

A:Cross-references: UNIPROT:Q05860; UNIPARC:UPI0000027927; EMBL:X53599; NID:952877; PIDN:

A:Query Match

Best Local Similarity 5.8%; Score 96.5; DB 2; Length 1468;

Matches 23; Conservative 2; Mismatches 24; Indels 13; Gaps 1;

Db 17 PPADPTGVDRDGPRIQVYIFSNDDVEYPPQPGGGLPDGDDGPPPPQPPPP 76

916 PPP 962

Qy 77 RL 78

Db 963 GL 964

RESULT 10

A:Accession: A53327

hypothetical protein BMEI0603 (imported) - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: A53327

R:DelVecchio, V.G.; Kapural, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

M.; Masur, M.; Goldstein, E.; Selkov, E.; Elker, P.H.; Hagluer, S.; O'Callaghan, D.; Letess,

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: A53327

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-189 <KUR>

A:Cross-references: UNIPROT:Q8Y142; UNIPARC:UPI0000057D45; GB:AE008917; PIDN:AL51784.1;

A:Experimental source: strain 16M

A:Genetics:

A:Gene: BMEI0603

A:Map position: 1

Query Match

Best Local Similarity 5.7%; Score 95.5; DB 2; Length 189;

Matches 47; Conservative 26; Mismatches 56; Indels 53; Gaps 10;

Db 166 VVNDLRYRPLSSA---VVRNALAH-----VGAKEKELSWRNSESFA 205

1 MODDLRIYDILAGELRGVIRKVLAEAKAGLPGNHFFITPLTGAPGVIRISRLKEXP 60

Qy 206 A-----WCRVGRRE-FKIGELRIGKOPRYLOIOLSAOR-----SITLFFOSL 247

Db 61 EGMWTLVLOHOFMAMHVTDLFEVGL--LSFGDIPKLTIPPSAIRGFDPSVNELEF--- 115

Qy 248 EDLIMEKRRNDQIGRAVLOELATHLPAPEEGSDNVART--TPPGRRPAP--SEED 304

Db 116 DVAVLDPASDNDGNSIPFEWA-----PEKADRPKAKTRKPAAGKESASTDSEOD 168

Qy 305 GE 306

Db 169 GE 170

RESULT 11

E29149

proline-rich protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 05-Oct-2004

C:Accession: E29149

R:Clements, S.; Menasho, H.; Carlson, D.M.

J. Biol. Chem. 260, 13471-13477, 1985

A:Title: Novel multigene families encoding highly repetitive peptide sequences. Sequence

A:Reference number: A92501; MUID:86033799; PMID:3840480

A:Contents: Clone pUM125

A:Accession: E29149

A:Molecule type: mRNA

A:Residues: 1-301 <CLS>

A:Cross-references: UNIPROT:Q62105; UNIPARC:UPI0000028BD2; GB:M11897; NID:G200540; PIDN:

C:Superfamily: Proline-rich peptide P-8

Query Match

Best Local Similarity 5.7%; Score 95.5; DB 2; Length 301;

Matches 32; Conservative 6; Mismatches 27; Indels 29; Gaps 6;

Db 1 MGNQVEXLTHLSYKE-----VPTADPTGVDRDGPRIQVYIFSNDDVEY-----PO--- 48

35 MGEQSGYGHHPKPPPGMPPRPSSGENDGDG-----SEEDVNGPERRPPQHP 87

Qy 49 -----PPPGGPGGGLPDGDD--GPPPPQPP 73

Db 88 HSGHHHGPPPG--DAQPPRPNGOQPPPPPPPP 120

RESULT 12

C87719

protein R119.6 (imported) - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: C87719

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.usci.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_e

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: C87719

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-549 <STO>

A:Cross-references: UNIPROT:Q61707; UNIPARC:UPI00000790D5; GB:chr_I; PIDN:AA616427.1; P

C:Genetics:

A:Gene: R119.6

A:Map position: 1

Query Match

Best Local Similarity 5.7%; Score 95.5; DB 2; Length 549;

Matches 72; Conservative 37; Mismatches 116; Indels 125; Gaps 15;

Db 18 TADPTGVDRDGPRIQVYIFSNDDVEYPPQ----- 50

210 TASVGVNPPPG-----YVFNNGRTPPPPPPPPPPPPPPPPPPPPPPPPPP 263

Qy 51 -----PPGPPGGLPDGDDP-----PPPOPO-----PY 74

Db 264 VGGGHHMVSAGARPPIRMPG--GGSPGGLGCPVAGPMGHQVQVNHPPPPPPPOIOQHNA 322

Qy 75 DPLHEVECSVFRDECTVQKSFAPGSAALSTYTPENILNK-----CK-PGLVEFV 125

Db 323 PVEWEVEENLQPTAAATATROYRPGSLKSTLKEDVILNRTKRMSSCSVERBALVAI 382

Qy 126 SOAQRPHAVVYGVNGVNVHLE---VINSFLTASQGRGRVNVNDLYRYKPLSSAVV 182

Db 383 SDVESHLELITLMAGVAENHVESLRIPENVVALDVYKROLFEEDDR----- 432

Qy 183 RNALAHVGAKEKELSWRNSESFAWCRVYKGEFKIGELRIGKOPRYLOIOLSAORSHTL 242

```

Db      243 AELFELVGPITLHNMKMAERREFEFLIEFCEHNGEPLTMVPOVYSQSIDLHRL----- 299
Qy      152 NSFPLTASQGRGRGVNDLYRYKPLSSSAVVRNALAHVAKERELSMRN-----SE 202
Db      300 --YICVRANKGQOYITKDKY-----MKNLCTEANDDLAE 331
Qy      203 SFAMCRVGRKREPKTIGGELRTGKOPRYRLOIQ--SAQRSHITLFEQSLIEDLIMEKRNDQIG 261
Db      332 SSAA-----GYQLRKAYORHLMLSECRGTGRNDEDEVAFAFKKORRKPAGG 380
Qy      262 RAAVLOELATHLHPAREEGDSNVARTPPRPPRA 297
Db      381 NAAAAAANAAGGQDGKQGHGAPGSGAPPPGPGCA 416

```

RESULT 6

H71332

hypothetical protein TP0369 - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C:Accession: H71332

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

rean, J.; Kralak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Ueberback, T.; McDo

nally, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; M0ID:98332770; PMID:965876

A:Accession: H71332

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-516 <COL>

A:Cross-references: UNIPROT:083384; UNIPARC:UPI00001396D0; GB:AE001216; GB:AE000520; NIT

C:Genetics:

A:Experimental source: strain Nichols

A:Gene: TP0369

```

Query Match      5.9%; Score 97.5; DB 2; Length 516;
Best Local Similarity 24.1%; Pred. No. 3.6; Mismatches 103; Indels 101; Gaps 17;
Matches 73; Conservative 26;

Qy      47 POPPGGPGGGGLPDGDDGPPPPPOPPYPRRLHEVSCSYVRDECIYOKSFAPGSAALST 106
Db      138 PEPPEPSP-SAGRPKGRAPVPPSPSPASPEPRAEVQAE-----PEPRADS- 181
Qy      107 YTPENLNKCKGDDLVFVSOA-----QYR-HWAVYVGNFQV-----VHLRL 148
Db      182 --PRAWVPEEPDEVPVRSRAVQLAVGQKLEVLVYGEQW-VYVGHETAQPGLRHYQRKL 238
Qy      149 EVINSPLTASQGRGRGVNDLYRYKPLSSSAVVRNALAHVAKERELSMRNSESFAWC 208
Db      199 EESHSLFTLSAE-REGDFVL-AFSYDFVPRGDFVSALAVKVPKKE----- 283
Qy      201 NGKREFKTIGGELRTGKOPRYRLOI-----QLSAQRSHITLFEQSLIEDLIMEKRND 258
Db      284 -----GLARVRAPEYRTVSSPPTVVSLSLPGTGT-----ERRAE 331
Qy      259 QI--AAVLOELATHL-HPAAREEGDSNVART--TPPP-----GRPPASSEE 302
Db      322 EGGTSG-SQRAAAHTGAPVQDQDTDTVAVAKAQHGTPRPDEKDRPTVGGGRDPVPSDAV 380
Qy      303 EDG 305
Db      381 AAG 383

```

RESULT 7

SS5345

protein-tyrosine-phosphatase (EC 3.1.3.48) PTP-PEST - mouse.

C:Species: Mus musculus (house mouse)

C:Date: 10-Oct-1995 #sequence_revision 01-Nov-1996 #text_change 05-Oct-2004

C:Accession: SS5345; S54261; 148666

R:Charest, A.; Wagner, J.; Shen, S.H.; Tremblay, M.L.

Biochem. J. 308, 425-432, 1995

A:Title: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein tyrosine

A:Reference number: 148666; M0ID:95289971; PMID:7772023

A:Accession: SS5345

A:Molecule type: DNA

A:Residues: 1-775 <CH>

A:Cross-references: UNIPROT:p35811; UNIPARC:UPI000017598B; EMBL:X86781; NID:9804997; PI

submitted to the EMBL Data Library, May 1995

A:Description: Murine protein tyrosine phosphatase-PEST, a stable cytosolic protein ty

A:Reference number: S54261

A:Accession: S54261

A:Molecule type: DNA

A:Residues: 1-127, 'M', 129-309, 'R', 311-775 <CH>

A:Cross-references: UNIPARC:UPI0000028ARD; EMBL:X86781; NID:9804997; PID:CAA60477.1; P

A:Experimental source: strain BALB/c

C:Genetics:

A:Map position: 5A3-B

C:Superfamily: protein-tyrosine phosphatase, non-receptor type 12; protein-tyrosine-pho

C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphata

F;58-282/Domain: protein-tyrosine-phosphatase homology <PT>

F;231/Active site: Cys (phosphocysteine intermediate) #status predicted

F;237/Binding site: substrate phosphate (Arg) #status predicted

```

Query Match      5.9%; Score 97.5; DB 2; Length 775;
Best Local Similarity 18.7%; Pred. No. 6.3;
Matches 53; Conservative 35; Mismatches 109; Indels 87; Gaps 9;

Qy      28 DCPRIGVSYFSDNDEDEVPQPPGPGGGGLPDGDP--PPQPOPYDRRLHEVC 83
Db      311 DQNEITGTWVSSISEKQDSSPPKPRTRSCLVEDAKKEILQPEPHVPEILTPSP 370
Qy      84 SYFYDECIYOKSFAPGSAALSTYTPENLNKCKGDDLVFVSOAQYPMHVVYVGNFQV 143
Db      371 SAFTVTTWQDS-----DRHFKPVLMHASP----- 397
Qy      144 HLRLVINSPLTASQGRGRGVNDLYRYKPLSSSAVVRNALAHVAKERELSMRNSE 202
Db      398 EGHPPADLNYSYKSAQDWQKS-----ESAIEHIDKLERLTSF----- 435
Qy      203 SFAMCRVGRKREPKTIGGELRTGKOPRYRLOIQLSAQRSHITLFEQSLIEDLIMEKR- 255
Db      436 -----EIK--KVPLQGPSPFDGNTLLNGHAIKIKSASSSVDRTSKQELS 481
Qy      256 -----RNDIGRAAVLOELATHLHPAREEGDSNVARTPPRPP 295
Db      482 AGALKVDVDSQNSCADCSAAHSHRAESSEESQNSHTPP--RP 523

```

RESULT 8

S24407

formin isoform IV - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: S24407

R:Jackson-Grusby, L.; Kuo, A.; Leder, P.

Genes Dev. 6, 29-37, 1992

A:Title: A variant limb deformity transcript expressed in the embryonic mouse limb defi

A:Reference number: S24407; M0ID:92112033; PMID:1339380

A:Accession: S24407

A:Molecule type: mRNA

A:Residues: 1-1206 <JAC>

A:Cross-references: UNIPROT:Q05859; UNIPARC:UPI0000027920; EMBL:X62379; NID:G51552; PID

```

Query Match      5.8%; Score 96.5; DB 2; Length 1206;
Best Local Similarity 37.1%; Pred. No. 14;
Matches 23; Conservative 2; Mismatches 24; Indels 13; Gaps 1;

```

```

Qy      17 PTAADPTGVDRDGPRIQGVSYFSDNDEDEVPQPPGPGGGGLPDGDP--PPQPOPYDR 76
Db      690 PPPPPPPPPSPGDP-----PPPPPPPPPLPVLLPLNSGSGPPPPPPPPPP 736
Qy      77 RL 78

```



```

: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Mathur, Brian
: APPLICANT: Fridde, Carl Johan
: TITLE OF INVENTION: No. 6919192el Human Kinases and Polynucleotides Encoding the S
: FILE REFERENCE: LEX-0332-USA
: CURRENT APPLICATION NUMBER: US/10/803,277
: CURRENT FILING DATE: 2004-03-18
: PRIOR APPLICATION NUMBER: US/10/116,326
: PRIOR FILING DATE: 2002-04-04
: PRIOR APPLICATION NUMBER: US 60/282,036
: PRIOR FILING DATE: 2001-04-06
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 6
: LENGTH: 703
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-803-277-6

Query Match
Best Local Similarity 21.6%; Score 122; DB 2; Length 703;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

77
-----RLHEVCSVFYRDECTYOKSFAPGSALSTYTPENLCKRPGD 120
461 PPSDGGGAGVGAAMRSRLINSIRNPSFGSPFRHKKQVPAEEMSLTPSSPELAKRS- 519
121 LVEFSAQVPMHAYVGFQVHLHRLVINSFLDASGRGRGVNDLRYRPLSSA 180
520 -----WFGNF--ISLDKEQIPLVLD-----KPLSS-- 544
181 VRNALAHVAKERELSWR--NSESPAAACRYGKREPKIGELRIKQPYRLQIQLSAQR 238
545 -IKADIVAHAFIPSLSHSVLSQTSFRA-----EYKASGCSVPQKPVRFQVDISSE 596
239 S-----HTLFQSLDLIMEKRRNDIGRAAVIQELATHLHPAPEEGD- 282
597 GPSPRRDGGGGGIVSTFTLISG--PSRRFKKVVETTQALLSTHDQPSVQALADE 653
283 SNVART-----TPPPGRP-PAPSGSEEDG 305
654 KNGAQTRPAGAPPRSLQPPRGPDPBELSSPRRG 687

RESULT 9
US-10-116-326-4
: Sequence 4, Application US/10116326
: Patent No. 6777545
: GENERAL INFORMATION:
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Mathur, Brian
: APPLICANT: Fridde, Carl Johan
: TITLE OF INVENTION: No. 6777545el Human Kinases and Polynucleotides Encoding the S
: FILE REFERENCE: LEX-0332-USA
: CURRENT APPLICATION NUMBER: US/10/116,326
: CURRENT FILING DATE: 2002-04-04
: PRIOR APPLICATION NUMBER: US 60/282,036
: PRIOR FILING DATE: 2001-04-06
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 4
: LENGTH: 762
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-116-326-4

Query Match
7.3%; Score 122; DB 2; Length 762;
Best Local Similarity 21.6%; Pred. No. 0.0014;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: ALRGN.062A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
;
US-08-808-303-8

Query Match
Best Local Similarity 32.2%; Score 136; DB 1; Length 164;
Matches 39; Conservative 15; Mismatches 45; Indels 22; Gaps 4;

QY 117 KPGDLVEFVSQAQYPRHAYVGNFQVHL-----HRLVINSFLTD 157
| | | | | : | : | : | : | : | : | : | : | : | : | : |
DB 9 KPGDLIE-IFRLGIEHMLYIGDGYVHLAPPSEYPGAGSGVFSVLSNAEYKGRLED 67
158 ASQGRGRGVNDL-YRKYPLSSSAVVRNALAHVAKER-ELSMRNSESFAAMCRGKREF 215
QY 68 VVGCCCYRNNSLDHEYOPRPEVVIISAKEMVGQMKYSIVSRNCEHFVTQLRYGKSRNC 127
DB 216 K 216
128 K 128

QY 216 K 216
DB 128 K 128

RESULT 5
US-08-996-533-8
; Sequence 8, Application US/08996533
; Patent No. 6294657
; GENERAL INFORMATION:
; APPLICANT: Nagpal, Sunil
; APPLICANT: Diserio, Daniel
; APPLICANT: Chandraratna, Roshantha
; TITLE OF INVENTION: RETINOID INDUCED GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,533
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,303
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: ALRGN.062A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
```

```

; TELEFAX: 714-760-9502
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
;
US-08-996-533-8

Query Match
Best Local Similarity 32.2%; Score 136; DB 2; Length 164;
Matches 39; Conservative 15; Mismatches 45; Indels 22; Gaps 4;

QY 117 KPGDLVEFVSQAQYPRHAYVGNFQVHL-----HRLVINSFLTD 157
| | | | | : | : | : | : | : | : | : | : | : | : | : |
DB 9 KPGDLIE-IFRLGIEHMLYIGDGYVHLAPPSEYPGAGSGVFSVLSNAEYKGRLED 67
158 ASQGRGRGVNDL-YRKYPLSSSAVVRNALAHVAKER-ELSMRNSESFAAMCRGKREF 215
QY 68 VVGCCCYRNNSLDHEYOPRPEVVIISAKEMVGQMKYSIVSRNCEHFVTQLRYGKSRNC 127
DB 216 K 216
128 K 128

QY 216 K 216
DB 128 K 128

RESULT 6
US-08-801-742-1
; Sequence 1, Application US/08801742
; Patent No. 6359123
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL H-REV107-LIKE
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,742
; FILING DATE: Filed Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0200 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-801-742-1
```

LENGTH: 778
TYPE: PRT
ORGANISM: Homo sapiens
US-10-003-690-2

Query Match 7.3%; Score 122; DB 2; Length 778;
Best Local Similarity 21.6%; Pred. No. 0.0017;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

50 PPGGDDGGLPDGDDGPPPO---PPPYD-----76
479 PSRCPGCGG---AGEOPPPSARSTPLPGPGSSSGTPLHSPHTPRASPTGTPTT 535
77 -----RLHEVCSEVYRDECIYOKSFAPGSAALSTYPTENLNKCKPGD 120
536 PPSPPGGVGGAAMRSRLNSIRNSFLGSPRRHRKMQVPTAEKMSLTPESSPELAKS- 594
121 LVEFVSQAOYPMHMAVYGNFQVHLHRLVINSFLTDSQGRGRGVNDLYRYKPLSSA 180
595 -----WFGNF--ISLDKEQIFLVLD-----KPLSS-- 619
181 VVRNALAHVAKERELSMR--NSESFAMCRYGKREFKIGELRIQKOPYRLQIOLSAOR 238
620 -IKADIVAHFLISPLSHSVLSQTSFRA-----EYKASGSPVQKVRQVDISSE 671
239 S-----HTLEFQSLIEDLIMEKRRNDIGRAAVLOELATHLHPAEBEED- 282
672 GPEPSRRDGGGGIYVTFLLISG---PSRRFRKRVETIQAILSTHDPVSQALADE 728
283 SNVART-----TPPGGRP-PAPSSSEEDG 305
729 KNGAOTRPAAGAPPRSLQPPGRDPPELSSPRRG 762

RESULT 13

US-10-803-277-2
Sequence 2, Application US/10803277

Patent No. 6919192
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Mathur, Brian
APPLICANT: Fridgde, Carl Johan
TITLE OF INVENTION: No. 6919192el Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0332-USA
CURRENT FILING DATE: 2004-03-18
PRIOR FILING DATE: 2002-04-04
PRIOR FILING DATE: 2001-04-06
PRIOR FILING DATE: 2001-04-06
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 778
TYPE: PRT
ORGANISM: homo sapiens
US-10-803-277-2

Query Match 7.3%; Score 122; DB 2; Length 778;
Best Local Similarity 21.6%; Pred. No. 0.0017;
Matches 72; Conservative 32; Mismatches 102; Indels 128; Gaps 14;

50 PPGGDDGGLPDGDDGPPPO---PPPYD-----76
479 PSRCPGCGG---AGEOPPPSARSTPLPGPGSSSGTPLHSPHTPRASPTGTPTT 535
77 -----RLHEVCSEVYRDECIYOKSFAPGSAALSTYPTENLNKCKPGD 120
536 PPSPPGGVGGAAMRSRLNSIRNSFLGSPRRHRKMQVPTAEKMSLTPESSPELAKS- 594
121 LVEFVSQAOYPMHMAVYGNFQVHLHRLVINSFLTDSQGRGRGVNDLYRYKPLSSA 180
595 -----WFGNF--ISLDKEQIFLVLD-----KPLSS-- 619

181 VVRNALAHVAKERELSMR--NSESFAMCRYGKREFKIGELRIQKOPYRLQIOLSAOR 238
620 -IKADIVAHFLISPLSHSVLSQTSFRA-----EYKASGSPVQKVRQVDISSE 671
239 S-----HTLEFQSLIEDLIMEKRRNDIGRAAVLOELATHLHPAEBEED- 282
672 GPEPSRRDGGGGIYVTFLLISG---PSRRFRKRVETIQAILSTHDPVSQALADE 728
283 SNVART-----TPPGGRP-PAPSSSEEDG 305
729 KNGAOTRPAAGAPPRSLQPPGRDPPELSSPRRG 762

RESULT 14

US-08-801-742-4
Sequence 4, Application US/08801742

Patent No. 6359123
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: A NOVEL H-REV107-LIKE
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,742
FILING DATE: Filed Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0200 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1709969
US-08-801-742-4

Query Match 7.0%; Score 116; DB 2; Length 160;
Best Local Similarity 31.9%; Pred. No. 0.00066;
Matches 38; Conservative 17; Mismatches 36; Indels 28; Gaps 6;

117 KPGDLVEFVSQAOYPMHMAVYGNFQVHL-----HRLVINSFLTDS-----159
7 KPGDLIE-IRPMYSHMAIVGDGVITHLAPSEIPGAGAASINSAITDAIYKKEILRD 65
160 -QGRGRGVNDLY--RYKPLSSSAVVRNALAHVAKERELSMR--NSESFAMCRYG 211
66 VAGKQKQVNNKHDKETTPLPNKTIQRAELVQ---QEVLYRLTSNCEHFVIELRYG 121

Db 67 -----GGSYPPPPSTGAAVAP-----PPGPAISRLPEKAVT 98
 QY 132 HW-----AAVYIGFQVHHLRLVINSFLTDAOGRRGVNDLYRKELSSAVVRNAL 186
 Db 99 FWTGRLAVLADINIPATVLLIGMLIOTLT-----KOEACVTDITOVNNOCAIOTPTG- 152
 QY 187 AHVAKKEIELSWRNSESFAAMCRVYKREPKIGELRIGK-----OPYRLQIQ 233
 Db 153 --IGMAFWFAWLMATAYLWM-NYGYRGATGSS--IGTYMKFKVISEATGPIGFGMS 207
 QY 234 IAGRSHTLEF-----QSLLEDLIM 252
 Db 20 VVROLAHFVDAVICIGIFLPLMDSKQTLADKIM 242

RESULT 3

T00051
 hypothetical protein KIAA0404 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
 C:Accession: T00051
 A:Shikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.;
 f:Submitted to the EMBL Data Library, October 1997
 A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The
 A:Reference number: Z14080
 A:Accession: T00051
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1956 <ISH>
 A:Cross-references: UNIPROT:Q63154; UNIPARC:UP100000726E0; EMBL:AB007864; NID:d1175320;
 A:Experimental source: brain, clone HG1236
 C:Genetics:
 A:Note: KIAA0404

Query Match 6.1%; Score 102; DB 2; Length 1956;
 Best Local Similarity 23.2%; Pred. No. 9.1; Indels 100; Gaps 14;
 Matches 72; Conservative 24; Mismatches 114;

QY 5 VEKLTSLSYKEVPTADPTGVDRDDGPRIGVSYFSDNDEDEVPQPPPGDGGGLPDGD 64
 Db 1312 LRELNOPSGLHPQAPISVYLFPGERSG-----PPSPVVG--PAGSL 1355
 QY 65 GPPPPQPPYDPLRHEVCSVFRDE-CIYOKSPAGSAAALSTYTPENLAKCKKPDLYB 123
 Db 1356 GSCSEKKE--DEEEEGDGDGLDSDDEFCLD--AGC-LGIPRDEPVTQLHPPI-- 1407
 QY 124 PVSQAQPYHMAVYVGFQVHHLRLVINSFLTDAOGRRGVNDLYRKELSSAVVR 183
 Db 1408 -----VVRDQYFSRPIGSTDLR 1425
 QY 184 NALAHVAKERELSWRNSESFAAMCRVYKREPKIGELRIGKQ- YRLQIOLSAQSHTL 242
 Db 1426 -APAHFPVSTRVLAR--EVLVHMLYGGDF-----GPHRGHARTGSGPRSSPS 1474
 QY 243 EFQSLLEDLIMEKRRNDQIGRA-AVLOELAT-----HLHPAPEGSDSVNATTPPRGR 294
 Db 1475 RCGPFRPQNSWMTQCGSGRHHVLMELQLSKVSFQHEVVPAPATG----- 1521
 QY 295 PPAPSEED 304
 Db 1522 PAAFSOELEB 1531

RESULT 4

T00328
 hypothetical protein KIAA0552 - human
 C:Species: Homo sapiens (man)
 C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C:Accession: T00328
 A:Shikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
 DNA Res. 5, 31-39, 1998
 A:Title: Prediction of the coding sequences of unidentified human genes. IX. The completed

A:Reference number: Z14086; MUID:96290545; PMID:9628581
 A:Accession: T00328
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-673 <NAG>
 A:Cross-references: UNIPROT:Q60299; UNIPARC:UP10000139A8C; EMBL:AB011124; NID:g3043627
 A:Experimental source: brain
 C:Genetics:
 A:Note: KIAA0552

Query Match 6.1%; Score 101.5; DB 2; Length 673;
 Best Local Similarity 22.6%; Pred. No. 2.4;
 Matches 69; Conservative 35; Mismatches 106; Indels 95; Gaps 12;

QY 56 GGGLPDGGDPPPPQ-OPYDRLHEVCSVFRDECTYOKSPAGSAAALSTYTP----- 110
 Db 304 GGGLPFAACSPSPSALIOLEERLEMEK-----DEVALARSLQSEAAVAQVLEEROK 358
 QY 111 -----NLNCKXQGDVEFVSQAQYHMAVYVGNFQVHHLRLVINSFLTDAOGRR 163
 Db 359 AMERELAEIRGCS--GKLQVARRAQ-----RAQGLQ 390
 QY 164 GRVNDLYRKELSSAVVRNALAHVAKERELSWRNSESFAAMCRVYKREK----- 215
 Db 391 LQVLRQDQKQLOSEA-----ALMRQREL-----EDKVAACQKEQADLPRIETK 439
 QY 216 ---KIGELRIGKQPYR-----LQIOLSAQSHLEFQ---SLEDLIM 252
 Db 440 WEVCQKAGISILKQOLKDSQADVQKSEIVGLRSQLEGGASJREKEQOLLSLDSFS 499
 QY 253 EKRRNDQIR-----AAVLOELATHLHPAPEEG-----DSNVARTTPPPGRPPAPSEED 304
 Db 500 SKQASLEEGELPAKCLPALTPVDPAPFOALATCESDEAKMRQAGVAAAASLVSD 559
 QY 305 GEAVA 309
 Db 560 GEAEA 564

RESULT 5

T29266
 hypothetical protein C01G8.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T29266
 R:Du, Z.; Gatlung, S.
 A:Submitted to the EMBL Data Library, November 1996
 A:Description: The sequence of C. elegans cosmid C01G8.
 A:Reference number: Z20597
 A:Accession: T29266
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-419 <DUZ>
 A:Cross-references: UNIPROT:P91019; UNIPARC:UP1000017B712; EMBL:U80439; PIDN:AA837646.
 A:Experimental source: strain Bristol N2; clone C01G8
 C:Genetics:
 A:Gene: C01G8.8
 A:Map position: 1
 A:Introns: 222/1; 285/3

Query Match 5.9%; Score 98; DB 2; Length 419;
 Best Local Similarity 19.9%; Pred. No. 2.5;
 Matches 67; Conservative 35; Mismatches 116; Indels 118; Gaps 12;

QY 14 KEVPTADPTGVDRDDG--PRIGVSYFSDNDE-----VEPQPPPGDGGGLPDGGG 65
 Db 147 KEGPNGNPATPSSSQPIPSPASSIAESLDDKSGTMAPQAPPQ-----QH 195
 QY 66 PPPPPQ-----PYDRLHEVCSVFRDECTYOKSPAGSAAALSTYTP-----NLNCK 115
 Db 196 PPPPQOQIMSPFPQAP-----SQATPSSSAASVAAPPTPKVSGVLSK 242
 QY 116 CKPGD-----LVEFVSQAQYHMAVYVGNFQVHHLRLV 151

